

USAID Program and Operations Assessment Report No. 4
A Performance-Based Budgeting System for the Agency for
International Development
By Allen Schick
June 1993
(PN-AXX-268)

[This on-line version does not contain any figures, charts, graphs, or tables. To receive a published copy of this report, contact DISC, 1500 Wilson Blvd., Suite 1010, Arlington, VA 22209-2404. Phone 703-351-4006; fax 703-351-4039; internet docorder@disc.mhs.compuserve.com]

Foreword

In response to the great interest in Performance-Based Budgeting which arose in the last few years in both the Executive Branch and the Congress, and especially to the stimulus of the Office of Management and Budget, government departments and agencies are studying ways to tie resource allocation to the likelihood of results. The Agency for International Development (A.I.D.) has committed itself to using a PBB system.

To assist the Agency to assess its current efforts and to bring the latest thinking to the attention of Agency managers, the Center for Development Information and Evaluation commissioned this study from Professor Allen Schick, of the University of Maryland, who is a nationally-recognized expert on Performance-Based Budgeting in government. Dr. Schick's paper is intended to present a PBB system appropriate for A.I.D. His analysis suggests, however, that there are limitations on the possible uses the Agency can make of PBB, limitations which may be somewhat more acute than those which affect other, domestically oriented, government entities.

The author makes a helpful distinction between allocating resources among countries and within countries. He argues that PBB should be used principally to allocate funds within rather than between countries. Moreover, even for making funding choices among sectors and projects within recipient countries, A.I.D. faces special obstacles to using PBB.

U.S. Government Experience

The underlying premise of PBB in government is that resources should be targeted toward specific results attributable to government action. These results should be different from what would have happened if government had not acted.

The first major section of the report, "Concepts and Experience," is an enlightening survey of attempts by the Federal government to use performance as a basis for the allocation of resources. It describes the techniques that have been tried since shortly after World War II and explains why each one of them either failed or

was abandoned. Performance budgeting, planning-programming-budgeting, management by objectives, and zero-based budgeting were all fashionable before the 1980s. Many of us experienced the application of these systems in A.I.D. Their failure to thrive is sobering and is probably explained, as suggested by the author, not by the bureaucracy's inability to devise and apply measures, but rather by the fact that inevitably too many factors other than performance actually affect budget decisions.

The 1990s are seeing a resurgence of interest in managing for performance. Techniques to measure performance are now in place in most Federal agencies, and a great deal of relevant information is being collected. With so much information now available, the problem is to determine what to do with it. In many agencies, the data collected is systematically reported to citizens or clients. Performance benchmarks expressed as quantified targets are announced in advance, so that progress can be observed. Performance audits go further still: auditors actually review performance statements made by agencies for reliability and accuracy.

Finally, PBB is on the scene, attempting to allocate resources against either expected outcomes or those actually produced. Most agencies, including A.I.D., can measure performance. But for A.I.D., it is not clear how closely allocation of resources should be tied to results. Too close or too rigid a linkage might diminish the role that judgment must play in making country and project allocations.

A.I.D.'s Experience

The author devotes the next section, "Allocating A.I.D. Resources," to reviewing A.I.D.'s 30 years of effort to find a rational approach to allocating resources. During this period, A.I.D. developed and relied mostly on programming techniques. Later (during the 1970s and 1980s) evaluation results were added to the mix of factors affecting allocation. Now A.I.D. is seeking to find a way to incorporate measures of performance.

Programming

Programming has taken many forms over the years as A.I.D. periodically tried to reinvigorate it and to strip away the layers of unnecessary requirements which inevitably accrete to every system. Its fundamental nature has, however, persisted. Programming creates a framework based on analysis of the circumstances and development requirements of the recipient country, usually undertaken by in-country staff. It then attempts to apply A.I.D.'s development assistance policies to the individual country case in order to come up with an appropriate level of assistance and funding levels for individual projects or programs.

The Country Development Strategy Statement (CDSS), developed in the late 1970s, was once considered the "ultimate" programming

document. Dr. Schick found, however, that many Missions were unable to provide the measurements the CDSS originally required, so it evolved into a document describing and explaining qualitative indicators of economic and political conditions. These indicators were not very helpful in allocating resources. At the same time, greater delegations of authority to the field for project approvals, the proliferation of congressional earmarks, and the extensive use of Economic Support Funds in the 1980s limited A.I.D.'s ability to allocate funds based on any results-oriented rationale.

Evaluation

A.I.D. has long believed that evaluating past experience would make it possible to make better decisions about the future. But the Agency, at least until recently, has not had a very good system of integrating the results of evaluations into resource allocation decisions. Indeed, for most of A.I.D.'s life, evaluation received relatively low priority, despite the fact that evaluations were expected to be performed regularly throughout the Agency, and many, in fact, were completed. The author notes various problems with using evaluation as a basis for budget decisions. First, evaluation findings are simply not included in whatever materials are used in planning and budgeting. It is also difficult to harmonize the programming function, which looks at countries, with the evaluation function, which typically looks at activities. At the same time, a mass of technical evaluation studies, however comprehensive, provides little guide to what is really happening to a country's level of development.

Moreover, evaluation of institution-building projects and policy reform, the kinds of efforts which have become standard fare since the 1980s, poses problems of its own. The lack of quantifiable data renders judgment difficult; the time it takes for effects to be felt exceeds the period during which the Agency can follow the activity; and the ultimate success or failure of activities like these depends on many factors other than A.I.D.'s performance.

Performance Management

To improve on the weaknesses of programming and evaluation as techniques for allocating resources, the Agency recently has undertaken three initiatives: a restructured programming system; a Program Performance Information System for Strategic Management (PRISM), and PBB. Taken together, these approaches assume that objectives can be arranged in hierarchical order, that objectives and results are measurable, that focus should be on those results contemplated when plans are made.

The new programming system is to rely on the measurable impacts of projects and programs on recipient countries' development as a means to allocate resources. The PRISM system is to tell A.I.D. managers whether the actual outcomes are in line with expectations. But however thorough the collection and

dissemination of information on performance is, some way must be found to ensure that the Agency will use that information for deciding what is to be funded and at what funding levels. For this reason, A.I.D. is trying to install a PBB system. Schick points out, however, that A.I.D.'s version of performance-based budgeting does not meet accepted definitions. The criteria now used for allocations are based on preexisting country conditions that are considered favorable for A.I.D. development efforts, rather than on the performance or results of A.I.D.'s programs. The A.I.D. system is evolving from versions first developed in the Latin America and the Caribbean Bureau (LAC) and the Africa Bureau (AFR). The LAC method weights indicators of economic and political conditions, suggesting how effective assistance is likely to be, and uses the result to establish country levels. AFR looks at similar indicators and adds factors about need and country size. Neither system, nor that now being developed by the Policy Directorate for agencywide use, relates resources to the planned or actual results of A.I.D. programs, which performance-based budgeting would normally be expected to do. On page 19, Dr. Schick states, "The A.I.D. approach is predicated on the expectation that the success or failure of assistance hinges on the policy environment within which development is taking place."

A System for A.I.D.

In the section, "A Performance-Based Management System for A.I.D.," Dr. Schick defines the conditions under which performance might be used to create an effective resource allocation system for A.I.D., makes specific recommendations, and explains the many reforms of the practices of both A.I.D. and the Congress that will be needed.

Allocations Among Countries

Dr. Schick argues that while indicators of country conditions and needs may be useful aids to making informed resource allocation decisions, using rigid statistical formulas to rank countries may be misleading; moreover it does not allow for numerous other factors that should be considered in allocations.

The measuring methods A.I.D. now uses are designed to permit countries to be given a score or ranking based on how well they may be expected to use assistance, judged on the basis of the prevailing economic and political conditions in each country (the "climate for investment") and, in Africa, the degree of need. While this grading system may carry the aura of objectivity, Schick points out that in any such a formula-based allocation, small changes in the weights assigned to various criteria can make a big difference in country rankings. Moreover, country rankings based on investment climate and need could be quite different from rankings based on project performance.

In addition to these concerns, Schick recognizes that foreign assistance is part of the foreign policy process and concludes that country allocations must be based on qualitative judgments

in order to factor U.S. political interests fully into account. Nevertheless, Schick argues that there is a role for using quantitative measures of development potential and need as aids to decision-makers in making country allocation judgments. He proposes a system, with several distinct steps, to facilitate the process. First, measures of development potential and need would be prepared and countries ranked or scored separately on each dimension. Then countries would be arrayed in a two dimensional matrix to provide a useful summary of which countries have the highest development potential and need and which have the lowest. Decision-makers could use this information, but not be bound by it, in making their allocation decisions.

Schick recommends that the country allocations made this way should be considered tentative and that they should only be finalized in a second stage on the basis of program and project objectives and targets. Missions would not be assured funding until they have specified the performance or results they promise to deliver with the requested funds. These tentative allocations would be modified to shift funds from lower to higher performing Missions and projects as performance data became available. However, the final allocations (like the tentative allocations) would be decided on the basis of many factors, enhanced by a quantitative analysis of performance, but not dictated by some mechanical rating.

Allocations Within Countries

As with allocations among countries, allocations within countries should be based to a significant degree on informed judgment, but here a PBB approach can provide much of the information needed for such judgments. Specific projects which are contributing to a country's performance must be identified. This is different from assessing a country's program performance, which A.I.D.'s performance measurement systems-especially PRISM-are now trying to do. Since the author argues that the outcome of specific projects can be measured in terms of each one's share in bolstering program performance, failure to undertake that measurement reflects a failure to justify the funds provided. Based on the findings of that measurement, and assuming that it can be made in a timely enough fashion to affect related decisions, Schick believes that funds can and should be obligated and de-obligated on that basis.

This kind of measurement is entirely consistent with an emphasis on results, not just plans or promises, as long as actual results are monitored and fed back into the programming process. Judgment enters, however, because it is typically impossible, even within a given country, to compare in some objective hierarchy the contribution to performance of each project in the portfolio. The objective should be to channel more resources where their use can be seen to induce an outcome different from and presumably better than what would have happened anyway if no funds had been provided. Conversely, activities not contributing to change should see budget cuts. Schick urges Missions to adopt a system of "variance" analysis, as it is practiced by the private sector,

but typically not by government. Since projects and programs turn out to be different from what was planned, we must know what those differences really were and adjust budgets accordingly. This approach does appear to have considerable usefulness in helping decide which activities to continue and which to terminate within a given country program. It would also be useful in identifying areas where project elements could be eliminated or activity levels reduced. Curtailing resources would be possible if, for example, it could be determined that certain elements were not causing the expected changes or that the same change could be produced at lower funding levels.

Dr. Schick has elaborated these ideas into a system to determine when project funding levels should be increased by measuring the additional results which can be attributed to increments in funding. This system does not always correspond well with the kinds of activities A.I.D. funds. In many cases, more money cannot be expected to increase output, either because quantifiable outputs are not the objective or because host country capacities are already taxed to the limit by the existing level of funding. But even if the particular construct proposed will not always work in the A.I.D. world, it is still helpful, at the outset, as an approach to deciding how to allocate the funds being provided for a particular country program.

Schick recognizes, however, that if funding allocations are going to be made on the basis of results, those supplying the information on results may be reluctant to do so for fear that it will be used against them. If this is what happens, the flow of accurate information will stop, as Missions will act defensively. Similarly, if performance information is ignored, then reporting will become perfunctory. The author posits a world where A.I.D./Washington managers will use the field-supplied information for wholesome dialogue and discussion, eschewing the requirement for overly frequent reporting of performance data to keep the focus on the long term, and neither demonstrate disinterest in the information nor use it to fill in a rigid formula leading to draconian funding changes. The reader may judge the likelihood of such a benign scenario.

Obstacles to Performance Management in A.I.D.

Professor Schick believes that there are cultural impediments to performance management in A.I.D. Unless they are overcome, he concludes that none of his proposals is likely to bring about the necessary reform. His prescriptions for cultural change are easily his most challenging ideas. They call for profound behavioral changes from A.I.D. and Congress, changes in attitudes and procedures that have evolved over time to cope with an ever-changing and often adversarial reality.

In his analysis of obstacles, Dr. Schick notes how difficult it is to link foreign assistance to development results causally and suggests that this problem played a part in the choice to focus on conditions for development instead of results. He realizes that development is both slow and uncertain, that numerous

exogenous factors influence development outcomes, and that results depend more often on actions by the recipient country than on those by A.I.D.

Congress has also played a counterproductive role, cutting resources while increasing the number of objectives and earmarking funds. This thin spread of resources, coupled with the inflexibility produced by detailed specifications on how and where A.I.D. funds are to be spent, has caused all the objectives to be underfunded and dramatically reduced the Agency's ability to respond to performance. Earmarks are a particular problem, since they provide funds in response to political clout without regard to results. Finally, the Congressional Presentation, which must now lay out explicit country and program levels so that Congress can examine every detail, telegraphs our intentions to recipient countries well before their performance can be measured and taken into account. Dr. Schick lays out a blueprint for a new relationship between A.I.D. and Congress, with Congress giving up its present micromanagement tools in favor of providing a clear definition of expectations and requiring information about results.

Schick also discusses internal Agency problems that he believes discourage a focus on results and inhibit PBB, such as lack of leadership to define common objectives and make substantive policy; overly decentralized programming and performance monitoring systems; and the proliferation of small low-impact projects. He laments A.I.D.'s overextended pipelines and mortgages, too, because they inhibit the flexibility needed to respond quickly to performance. He argues convincingly and makes sound recommendations for better management of slow-moving projects and those which are not meeting their targets. Despite the many hard choices it calls for, this study, I believe, illustrates effectively what it would take for A.I.D. to truly apply performance factors in allocating assistance funds. It thus sets a challenging agenda for action. How much of the agenda we are able to accomplish will govern our progress in moving toward Performance-Based Budgeting.

John R. Eriksson
Associate Assistant Administrator, Center for Development
Information and Evaluation, Directorate for Policy, June 1993

Summary

During the past 2 years, the Agency for International Development (A.I.D.) has moved to a new style of performance management that emphasizes the results and impact of development. Its initiatives have included a restructured programming process; PRISM (the Program Performance Information System for Strategic Management); increased investment in measuring, monitoring, and assessing performance; and performance-based budgeting (PBB). Although still in development, these systems are building on progress made in recent years by various bureaus and Missions.

These initiatives have come during a period in which the Agency

has undergone critical scrutiny by the General Accounting Office, the A.I.D. Inspector General, a Presidential Commission, and a joint Office of Management and Budget/A.I.D. SWAT team. One thrust of the resulting reports is that A.I.D. is a troubled agency that must develop the strategic capacity to establish and implement objectives, as well as the capacity to evaluate and redirect operations and resources to more productive and higher priority programs and activities.

The Design of a Performance-Based Budgeting System

A.I.D.'s determination to manage for results must be backed by two critical changes. One is to make results'actual and projected'the basis of program and resource decisions; the other is to transform A.I.D. operations so that the Agency has the inclination and capacity to take performance seriously. The first entails implementing some version of PBB; the second requires truly far-reaching changes in Agency behavior. The two reforms are interdependent, because without a transformation of A.I.D.'s operating culture, efforts to install PBB will be futile, and without established procedures for using performance data and assessments, the Agency will revert to its old ways.

A.I.D. has embraced PBB as a means of allocating resources on the basis of actual and expected results. PBB is a new name for an old idea that has been adorned with contemporary features. The basic idea is that government agencies should budget on the basis of outputs to be produced or results to be achieved, instead of on the basis of the previous year's spending patterns or the inputs to be acquired. Some versions of PBB could make the budget into a "contract for performance" in which incremental resources are closely tied to incremental results. In practice, however, the linkage of resources and results tends to be looser, ranging from budgets that estimate the unit cost of expected outputs to budgets that merely identify the work or activity to be carried out.

A.I.D.'s version of PBB emphasizes each assisted country's "climate for investment" rather than the specific impacts of development programs. The PBB system devised by the Policy Directorate (which draws on experience in the Africa and the Latin America and Caribbean Bureaus) measures each country's macroeconomic policies and assigns statistical weights to preselected sets of economic practices. This approach can be expanded to include political conditions, measures of need, and other variables affecting investment decisions.

A.I.D.'s "climate for investment" strategy is based on the notion that a country with favorable conditions will make more productive use of U.S. assistance than will a country with less favorable policies. But this approach does not meet accepted definitions of PBB. It confuses inputs for outputs and the conditions of development for the results of development. Performance should always denote the results ensuing from government action-not what exists in the absence of such action. Moreover, performance-based allocations should distribute

resources on the basis of projected or actual outcomes, not on the basis of conditions before assistance is provided.

PBB should be used principally to allocate funds within--not between countries. Although the notion of a statistical formula for parcelling funds among countries has considerable appeal, it would be more sensible to use performance criteria for allocating funds within countries and to rely more on informed judgment and less on statistical rankings for apportioning funds among countries. In any formula-based scheme, small changes in the weights assigned various criteria can make a big difference in country rankings. Rankings also change if they are adjusted for country size, measures of need, and other factors. It is misleading therefore to impute objectivity to rankings that are derived from subjective judgments on the relative value of criteria and weights. There is a useful niche in A.I.D. for formula-aided allocations, but this should be distinct from PBB. Economic and political criteria should serve as aids to judgment, not as rigid criteria for allocating funds among countries. The allocation of funds to a country is inherently a political decision. It entails judgments concerning U.S. interests, the effectiveness of previous assistance, the future course of development, each country's capacity to absorb external aid, the stability of a country's regime and durability of its commitment to open markets, and the contributions of other donors. These judgments cannot be boiled down into a simple formula.

Yet it would be inappropriate to disregard measurable characteristics of development in setting country assistance levels. The two most important characteristics are a country's investment climate, as measured by economic policies and political conditions, and its need for external aid, as measured by the prevalence of poverty and other adverse conditions. A.I.D. should take account of these characteristics by developing two measures--development potential and development need--each scored separately. Development potential would be measured by criteria similar to those currently used by A.I.D. and should include both economic and political factors. A uniform weighing scheme would be applied throughout the Agency. Development need would be measured in terms of comparative data on infant mortality, per capita income, and access to education or other services. Each country should then be ranked on the basis of its "potential" and "need" scores. These rankings should be arrayed in a matrix, from which A.I.D. officials then decide on country allocations, taking the size of the country into consideration.

Country allocations should be finalized on the basis of performance objectives. Allocations to countries cannot be decided independently of program and project decisions. Regardless of how favorably a country ranks in development potential and need, A.I.D. must still determine whether assistance can be productively used. To make this determination, allocations based on potential and need should be treated as tentative, with final allocations based on each country's projected performance, as specified in the programming process and project plans. A.I.D. must be willing to shift funds among

countries and from lower to higher performing Missions and projects.

Project plans and proposals should detail expected performance. The process of finalizing country allocations requires a sharp focus on what projects are contributing to performance. Projects are A.I.D.'s means of delivering results. Without a strong connection between projects and development, there would be little justification for U.S. assistance. Accordingly, A.I.D. should program, budget, and implement each project in light of its contribution to the country's progress. A project-oriented PBB system will provide bureaus and Missions data on how different project mixes and funding levels affect performance. This information will facilitate trade-offs between similar projects in the same country, or between projects in different countries.

Performance should be measured in terms of change from baseline or current levels. Performance should always refer to change. To say that a program is performing should mean that it is producing different outcomes than would prevail in its absence. PBB allocates resources on the basis of the changes to be purchased with government funds. Change, especially critical in development, has meaning only in the context of differences between a country's past and future condition. The measurement of change should focus A.I.D. on its overriding purpose; spurring improvement in the condition of developing countries.

A.I.D. should inform Missions on how performance measures will be used in programming and budgeting. PBB depends on trust between suppliers and users of performance data. The surest way to impair this trust is to keep producers of the data in the dark about how the measures will be used, by whom, and for what purposes. The best way to build trust is to encourage those who generate performance data-especially the Missions-to use the information in their own work: when they plan programs, justify their budgets, communicate with the bureaus and other headquarter units, review projects, and assess results.

A.I.D. should protect long-term development objectives against short-term fluctuations in performance. Performance measurement should not be applied in ways that shorten A.I.D.'s time perspective. Except when shifts are truly dramatic, U.S. assistance should focus on long-term objectives, typically 5 to 10 years in the future. Toward this end, quarterly performance reports should concentrate on key outputs, whereas detailed impact data should be reported annually. Performance reports should regularly assess the impacts of recent results on long-term objectives, including output targets and projected outcomes.

A.I.D. Missions should systematically analyze variances from expected performance and take appropriate corrective steps. The key to performance is not whether results have been promised but whether they have been achieved. PBB must not only be a method for preparing the A.I.D. budget, it must also be the basis on

which the budget is implemented. This approach requires the systematic comparison of results to budgeted objectives, the assessment of any variances, and adjustments in the budget and program.

Changing A.I.D.'s Operating Culture

Numerous characteristics of A.I.D. operations inhibit sustained attention to results. The pace of development is uncertain, slow, and halting and is typically measured in decades. Performance depends more on what assisted countries do than on A.I.D. action. The focus on performance is also clouded by Congress, which has steadily increased the number of objectives the Agency must fulfill without increasing its resources. Some of the damage to A.I.D.'s capacity to manage for results is self-inflicted. The Agency maintains an extended pipeline of projects; for many, initiation or completion is years away, if ever. Virtually all new funds are mortgaged to ongoing projects, and new projects often must wait for funding that may be years away. The mortgage causes programming to be out of synch with the budget cycle. Performance objectives from headquarters come too late to influence projects already underway and too early for new projects that will take years to move from the drawing board to fruition. Field managers begin their tour inheriting old projects programmed years earlier and end their tour by bequeathing planned initiatives to the next shift.

Congress should clarify A.I.D.'s development objectives and specify criteria for assessing progress. An essential task in managing for results is to remove accumulated obstacles. Congressional action is a prerequisite for Agency self-improvement. Without congressional action, A.I.D. will have neither the license nor the backbone to reform itself. The hodgepodge of 40 or so legislated objectives is anything but a roadmap for the Agency. These objectives give lip service to all purposes and priority to none, ensuring that A.I.D. will fall short of the mark, no matter how hard it tries. Congress should restate A.I.D.'s objectives and rethink the aims of foreign assistance, of the organization and operations of A.I.D., and of its own role. Congress should state the A.I.D. objectives precisely so they can serve as benchmarks for measuring results. The Congressional Presentation should have less country and project detail and should focus on performance. The Congressional Presentation enables Congress to dictate spending priorities and monitor A.I.D. operations. But there is a fundamental contradiction between detailed country lists and genuine consideration of results. Discussing performance is of little use when Missions and host countries know in advance what will be funded. Country-Agency negotiations must be at the heart of A.I.D.'s orientation to performance. These negotiations should not be prejudiced or encumbered by preset funding commitments. Accordingly, the Congressional Presentation should be overhauled to stress objectives and to preserve the Agency's negotiating and operating flexibility. Congress should not give A.I.D. a blank check, but it should accept less spending detail in exchange for more information on program impacts. The Congressional

Presentation should systematically inform Congress on achieved and expected performance.

Congress should curtail earmarks. Each year, Congress prescribes where the bulk of foreign assistance should go. These earmarks now consume a much higher percentage of A.I.D. funds than in the past. Earmarks send the wrong message that performance does not matter because funds are given to those with political influence. The Africa Bureau has demonstrated that relief from earmarks can lead to a coherent and focused development strategy. Earmarks will not disappear, but their progressive expansion should be reversed.

A.I.D. management must provide clearer, effective policy guidance. A.I.D. is a highly decentralized agency--a characteristic that leads to uneven implementation of Agencywide policy. Although decentralization is essential for operating in 90 countries, A.I.D. needs strong leadership, strategic vision, and the ability to mobilize resources and change course. Central guidance is especially urgent to move the Agency toward common objectives and to ensure compliance with congressional requirements. For senior managers to lead, the Policy Directorate must be involved in making substantive policy, not only in developing systems design and procedures. When top management leads, it has the capacity to redirect projects, to reprogram funds, and to take other steps that keep bureaus and Missions attuned to its agenda.

Projects should be consolidated into fewer activities with significant impacts on development. A.I.D.'s portfolio of approximately 2,000 ongoing and new activities is difficult to manage and diverts attention from performance. The plethora of projects means that particular activities tend to be underfunded, typically take years to complete, are difficult to evaluate, have insufficient impacts on development, and often have inadequate internal controls. All these eat into the Agency's capacity to keep its eye on results. To improve project definition and management, A.I.D. should curtail the number of projects. Each approved project should be of sufficient size and scope to have a critical effect on development. Each should make a measurable difference.

A.I.D. should shorten its pipeline by deobligating troubled projects and enforcing realistic timetables. A.I.D. has a large pipeline, one-and-a-half times annual appropriations. Although pipelines are unavoidable and sometimes advantageous for project success, A.I.D. should be sure that its pipelines are not caused by troubled projects whose implementation problems are impeding the completion of project objectives. In such cases, A.I.D. should make more vigorous use of its authority to deobligate and reobligate funds to free resources for use in successful projects. Headquarters must take the initiative in this regard because Missions may be constrained from acting by their desire to maintain stable relations with the host country.

Glossary

A.I.D. Agency for International Development
 API Assessment of Program Impact
 CAP Country-Assisted Program
 CDIE Center for Development Information and Evaluation
 CDSS Country Development Strategy Statement
 DAP Development Assistance Program
 GAO General Accounting Office
 IPA Indicative Planning Allocation
 MBO management by objectives
 LAC Bureau for Latin America and the Caribbean
 OECD Organization for Economic Cooperation and Development
 OMB Office of Management and Budget
 PBB performance-based budgeting
 PPB planning-performance-budgeting
 PRISM Program Performance Information System for Strategic Management
 ZBB zero-based budgeting

1. Introduction

Performance-based budgeting (PBB) is a new name for an old idea adorned with contemporary features. The idea is that governments and their agencies should allocate resources on the basis of outputs to be produced or results to be achieved, instead of on the basis of the previous year's expenditures or the inputs to be purchased. In pure form, PBB makes the budget into a "contract for performance," which closely links incremental resources to incremental results. In practice, however, the linkage of resources to results tends to be looser and ranges from budgets that specify the unit cost of expected outputs to budgets that merely call attention to the work or activity to be carried out. A number of developments have fueled the current push for PBB.

First, the Senate passed legislation in 1992 requiring that the annual budget establish measurable performance standards and goals and compare program results with the budgeted standards. This bill has spurred interest in performance-based budgets.

Second, the Chief Financial Officers Act of 1990 (Public Law 101-576) provides for the integration of budgeting with other financial management operations through systematic measurement of performance, annual progress reports, and auditable financial statements. Third, pursuant to this legislation, on February 5, 1992, the Office of Management and Budget (OMB) issued guidelines for developing performance measures for inclusion in annual financial statements. OMB will also establish 14 interagency teams to devise common performance measures for use by agencies engaged in similar activities. Fourth, the Governmental Accounting Standards Board, which has cognizance of State and local but not Federal practices, has issued a series of reports examining the feasibility of establishing service effort and accomplishment indicators for various government functions (Hatry et al. 1990). The Board concluded that current performance measurements are sufficiently developed to warrant the

presentation of such measures in financial reports or as separate documents. Fifth, the spread of total quality management practices in Federal agencies has stimulated interest in measuring citizen and customer perceptions of government performance. Sixth, Sunnyvale's successful and much publicized implementation of PBB has demonstrated the feasibility of this approach in allocating public resources. {Footnote 1} Seventh, in the past few years, it has become common to publish performance data on various critical government functions, such as the mortality rates in hospitals and grade-by-grade test scores of public schools. The emphasis on performance measurement has been spurred by fiscal stress, which has pressured governments to seek more value for public money, and by citizen discontent, which has pressured governments to disclose what they are doing with public funds.

These pressures have strongly affected the Agency for International Development (A.I.D.), which has come under intense scrutiny in recent years from Congress, the General Accounting Office, OMB, and other external sources. A.I.D. has responded by undertaking a major reorganization, establishing a Program Performance Information System for Strategic Management (PRISM), and committing itself to implementing a PBB system.

Despite the impetus for PBB, allocation of resources based on actual or promised results is neither simple nor obvious. It is not simple because of numerous organizational impediments; it is not obvious because difficult decisions must be made concerning what type of system to introduce. In moving toward performance-based allocations, A.I.D. faces special problems that will be far more difficult to surmount than those facing most other Federal agencies. Foreign assistance (and foreign policy generally) is not the most favorable arena for managing on the basis of quantifiable measures of performance. If A.I.D.'s special circumstances are ignored, PBB will go the way of other management innovations, starting with much fanfare but ending with few traces left behind. But if PBB is carried out as part of a careful strategy to remake A.I.D.'s operating system, it may succeed where other management innovations have failed.

This report is based on the belief that a performance-based system can be put in place more easily by recognizing (1) relevant past and current efforts to allocate Federal resources on a performance basis and (2) A.I.D. efforts over the years to emphasize performance in its programming and budgeting systems. Accordingly, Section 2 of this report surveys past innovations and explains why they generally fell short of expectations. It also briefly discusses some continuing developments that may affect the progress of PBB. Section 3 concentrates on what A.I.D. has done in managing for performance. It seeks to explain why the Agency has faced this issue repeatedly over the past three decades. Section 4 presents a series of recommendations for implementing a PBB system in A.I.D., stressing the conditions necessary for success and the organizational strategies and changes needed for the Agency to implement this reform.

2. Performance-Based Budgeting: Concepts and Experience

Since its emergence around the turn of the century, public budgeting has been concerned with what governments are getting for their money. During their formative years, however, most government budget systems concentrated on controlling expenditures by budgeting for the cost of personnel, supplies, travel, and other line items. The first nationwide effort to orient budgeting to performance was initiated by the 1949 report of the Hoover Commission (officially the U.S. Commission on Organization of the Executive Branch of the Government), which argued that "the all-important thing in budgeting is the work or service to be accomplished, and what that work or service will cost." In line with this view, the Commission urged that "the whole budgetary concept of the Federal Government should be refashioned by the adoption of a budget based on functions, activities, and projects; this we designate a 'performance budget'" (p. 8).

This early version of performance budgeting differed in one critical way from contemporary PBB. Its focus was on the work or activity being performed, in contrast to today's emphasis on objectives and results. But some versions of early performance budgeting bear a strong resemblance to today's concept. For example, the following is a working definition of performance budgeting devised 40 years ago in the U.S. Bureau of the Budget (the predecessor to OMB):

A performance budget is one which presents the purposes and objectives for which funds are required, the costs of the program proposed for achieving those objectives, and quantitative data measuring the accomplishments and work performed under each program (Burkhead 1956, 142).

Looking back at performance budgeting in the Federal Government, one can see that few agencies met this ambitious definition. During the 1950s, performance budgeting came to mean, as Jesse Burkhead (1956, 139) pointed out, "something different in every jurisdiction which puts it into operation." For many Federal agencies, it simply meant a reclassification of line-item budget accounts into functions and activities. Some agencies went a bit further and divided their accounts into measurable work units. And to encourage this development, the Bureau of the Budget published a series of guidebooks on work measurement. {Footnote 2} The work measurement approach had its fullest application in State and local governments that broke down their budgets into measurable work or cost units. For example, some police departments added data to their budgets on the volume of arrests, traffic citations, investigations, and other law enforcement activities. In a relatively small number of jurisdictions, the budget showed the unit cost of work to be performed. Few, however, actually budgeted resources based on the cost per unit of work. To do so would have required much more sophisticated cost accounting procedures than were (or are now) available in the public sector (such as distinctions between fixed and variable costs and between marginal and average costs).

Performance budgeting did not fail in the Federal Government; it simply faded away. Before performance budgeting could mature, it was eclipsed by program budgeting and planning-programming-budgeting (PPB) systems that promised a more fundamental reorientation. Within a few years after PPB emerged, hardly anyone spoke about performance budgeting, although many local governments retained work and output measures in their budgets.

One can draw several relevant lessons for contemporary budgetary developments from the experience of performance budgeting: Traditional accounting and reporting systems designed to guard against improper expenditure were inadequate for broader managerial tasks. The primitive state of cost accounting in the public sector in particular impeded cost analysis and allocation. Performance budgeting was a reaction to line-item controls and other constraints on managerial discretion. But the loosening of these controls was not accompanied by the establishment of performance-based accountability systems. Whether managers still felt constrained or were otherwise unable to exploit the new opportunities, they did not act as if they had discretion to budget for results.

The linkage of incremental resources to incremental results was hardly explored. The basic question of what a marginal increase in spending would buy in added output could not be confidently answered. Even when productivity or other output data were available, they were used merely as supporting information, not as the basis for deciding the budget. In explaining the chronically weak link of resources and results, it is important to note that governments, unlike many businesses, do not have variable budgets that automatically adjust to changes in current output measures, such as the volume of production or sales. Variable budgets are de facto performance budgets; fixed budgets are not.

Performance budgeting was not the only movement to orient governments to focus on outputs and results. Three subsequent reforms had elements of performance measurement. The first reform was planning-programming-budgeting (PPB) introduced in the Defense Department in the early 1960s and made governmentwide by Presidential order in 1965. The second was management by objectives (MBO), introduced by Richard Nixon in 1972 and continued by Gerald Ford when he took office. And the third was zero-base budgeting (ZBB) initiated by Jimmy Carter in 1977. No comparable innovations were tried in the 1980s, although other significant modifications were made in Federal budget practices during the decade.

Planning-Programming-Budgeting Systems

As the name suggests, PPB was designed as a three-stage process that began with planning multiyear objectives, continued with analysis and selection of programs to accomplish intended objectives, and culminated in annual budgets that allocate

resources to the selected programs. Thus, according to PPB, the budget would purchase the objectives planned by the government. Performance measurement played a prominent part in PPB through analysis of the cost-effectiveness of alternative means of pursuing approved objectives. In its heyday, PPB promised governments that they would gain the capacity to select programs that would yield the greatest increment in results for the dollars spent. {Footnote 3}

Although PPB shared much of the terminology of performance budgeting, some key performance budgeting concepts have different meanings. For example, in contrast to performance budgeting, which treated objectives as fixed targets, performance budgeting regarded objectives as options open to analysis and reformulation. In PPB, every objective was tentative and could be modified as new options or analysis became available. This broader view of objectives made it more difficult to measure progress and blurred the line between policy analysis and political action.

To a far greater degree than occurred in other postwar budget innovations, PPB was institutionalized in Federal agencies through the establishment of analytic staffs and the recruitment of highly trained analysts. As a result, when agencies discarded PPB, as most did in the 1970s, legions of specialists remained to conduct analytic studies and evaluations, even though most of the system's formal trappings, such as multiyear plans and program classifications, were jettisoned. (A few agencies, most notably the Defense Department, have retained PPB-type procedures.) Although analytic work continued, PPB failed in its principal aim of remaking Federal budgeting into a means of defining and pursuing long-term objectives.

It is generally agreed that PPB was a cumbersome system that failed for two reasons: (1) It imposed enormous analytic and data gathering burdens on agencies, and (2) it overestimated the prominence of analytic findings in allocating resources and was naive about the political context in which budgets are made. A January 1992 report prepared by the Organization for Economic Cooperation and Development (OECD) concluded that the failure of PPB led OECD member countries "to recognize that information obtained from performance measurement, as from policy and program evaluation, constitutes only one of the many elements involved in a policy decision" (Public Management Service 1992, 10). This conclusion surely applies to PBB as well.

Management by Objectives

Early in his second term, Richard Nixon launched MBO. Modelled after private-sector MBO systems then in vogue, this reform reverted to the idea of fixed targets rather than iterative objectives. But in contrast to performance budgeting, MBO was not confined to the budget. It was conceived of as a broad management strategy for energizing Federal agencies to establish and achieve specific objectives. Nixon trumpeted this strategy in his 1975 budget submission to Congress in February 1974:

To help keep a constant focus on program results, I have asked each major department and agency to work with me in developing a set of specific objectives ...These objectives will not simply be identified and then filed away and forgotten. Specific results are to be achieved by specific deadlines. These commitments will be reviewed continually and will guide day-to-day operations until the objectives are met (Budget of the U.S. Government 1974, 20).

As applied in the Federal Government, MBO meant each agency nominated a number of objectives that were reviewed and approved by the President. More than 200 such objectives were selected, and a process was established in each agency for monitoring progress, including setting specific milestones and periodic meetings between Presidential and agency staffs. However, few of the objectives established in 1973 and 1974 were of genuine Presidential interest. Many objectives pertained to procedural matters, such as deadlines for submitting reports, rather than to program results. {Footnote 4} It did not take long for interest in the new process to wane, as senior staff passed the work on to subordinates or routinely canceled scheduled 'milestone' meetings because more pressing issues demanded attention. The MBO exercise proved that "getting action on objectives is a very different matter from getting objectives listed on paper" (Rose 1977). Although MBO disappeared leaving few traces, it did provide one important lesson. The pursuit of objectives (and other management innovations) is not likely to succeed unless it is linked to ongoing activities that cannot be ignored, such as the annual budget process. Performance measures alone will either wither or be stored away.

Zero-Base Budgeting

Almost immediately after taking office in January 1977, Jimmy Carter ordered the implementation of ZBB in all Federal agencies. In name and theory, ZBB called for constructing the budget without reference to past allocations or commitments. Its avowed purpose was to rid budgeting of the incremental practices that had dominated resource allocations for many decades. {Footnote 5} In application, however, ZBB was a process that facilitated the marginal analysis of incremental budget performance. But this feature of ZBB never realized its potential, and reform was frustrated by misunderstanding, excessive paperwork, and the difficulty of building budgets from a zero base (U.S. Senate, Committee on Government Operations 1977).

The linkage of incremental resources with incremental performance in ZBB resulted from the manner in which administrative units constructed their zero-base budgets. Each unit was required to array its budget request in a series of decision packages, typically beginning with a "minimum" package below the previous year's spending level, then adding a "base" package that would bring spending up to the previous level, and finally adding one or more "enhancement" packages. Each decision package set forth the marginal resources that would raise spending to the next

higher level as well as the additional performance that would accrue from the incremental expenditure. In this way, the decision packages were built as increments on one another, adding up to a unit's total budget request (i.e., the full set of packages).

With the change of administration in 1981, ZBB vanished, but even when the White House promoted it, ZBB did not effectively match incremental resources and results.

Lessons From the Four Reforms

The review of the four innovations illustrate the difficulty of making performance measurement a determining factor in allocating government resources. The problem is not that performance cannot be measured; it can, but that these measures cannot always form the dominant basis for making budgeting decisions. Too many other factors intrude and obscure matters of performance when budget decisions are made.

There is no sure way to give preference to results and outputs in public management. Such attempts have failed regardless of the approach taken. Performance budgeting and MBO relied on fixed targets. PPB sought broader and more open-ended objectives, and ZBB expressly linked performance to spending levels, in contrast to MBO, which focused on overall managerial capacity. Some reforms (e.g., PPB and ZBB) were framed in the annual budget cycle, but PPB was ambitiously stretched to a multiyear context. Yet failure does not mean that American Governments have been indifferent to what they are doing with public funds. Their repeated efforts to build performance measures into budget procedures attest to the attractiveness of performance as a guideline for allocating resources.

As a result of past reforms, Federal agencies are stocked with many more indicators of performance than is generally realized. Perhaps the measures are not used effectively, but the solution is not to measure more things; it is to devise means of using the data already available.

Contemporary Developments

The 1980s were an inactive period in budget innovation. During the decade, huge deficits and protracted conflict between the President and Congress over spending and revenue policies dominated the budgeting process. It was not a decade of institution building, either in budgeting or in other arenas of government, although important initiatives in financial management laid the groundwork for the Chief Financial Officers Act of 1990.

The early years of the 1990s have given some indication that the decade will be more open to institutional changes, especially with such developments as PPB. The idea of performance is rife in contemporary innovations, appearing in various forms that differ in the extent to which performance measures are used in

allocating resources, informing citizens, and holding governments accountable for results. Five gradations of managing for performance can be identified, ranging from measuring outputs to auditing results.

Performance Measurement

Performance measurement is the simplest and most extensive application of managing for performance. Over the past 30 years, much progress has been made in distinguishing among different types of measures and gathering data on performance. A recent survey of more than 100 Federal agencies (accounting for more than 85 percent of all Federal employees and more than 90 percent of total outlays) indicates widespread use of performance measures. All of the agencies surveyed claim to measure final outputs or products, 93 percent have work or activity measures, 91 percent measure the timeliness of service, 83 percent have internal measures of quality, 71 percent collect data on outcomes, and 63 percent assess customer satisfaction (U.S. General Accounting Office 1992). It is now common to distinguish between measures of efficiency and effectiveness, as well as between measures of outputs and outcomes. These measures are prerequisites for more advanced use of performance data. While it is often recommended that agencies collect more data on performance, the emphasis now is more on what to do with measures already on hand.

Performance Reporting

Performance reporting is the systematic communication of data on results to citizens or clients. Systematic means that key indicators are selected in advance and reports on indicators are issued regularly. However, performance reporting goes beyond the standard material often found in government or agency annual reports. It looks at results from the perspective of those who pay for or are affected by public services and seeks to determine how well their interests are being served. Performance reports sometimes concentrate on a specific policy area, such as school performance as measured by standard test scores or hospital performance as measured by the variance between expected and actual mortality rates. An emerging emphasis in performance reporting is the issuance of a comprehensive report on the quality of government. In March 1992, for example, the City of Portland, Oregon issued its first annual performance report containing data on spending, workload, and results for the city's six most important public services. It also presented the results of a citizen survey and compared Portland's performance with results in other cities (Office of the City Auditor 1992).

Performance Benchmarks

Performance benchmarks take the performance management process one step further by specifying in advance the results expected. Benchmarks can be either short-term targets tied to current actions or long-term visions of where the government or community should be heading. Unlike the vision statements often found in

public documents, benchmarks are quantified, so that progress in achieving them can be monitored. Oregon has probably taken this development further than any other State. In 1991 the Oregon Progress Board (a State agency) issued Oregon Benchmarks: Setting Measurable Standards for Progress, which measures progress at six fixed intervals: 1970 and 1980 provide a historical backdrop, 1990 is the baseline year for measuring future progress, and 1995, 2000, and 2010 are the years for which benchmarks have been fixed for the future (Oregon Progress Board 1991).

Thus far, one feature of benchmarking found in the private sector does not appear to have made much headway in U.S. Government agencies. Some firms define benchmarks as the best practices (or results) found in competing firms or in other organizations facing similar situations. They then strive to match or surpass these benchmarks. In the public sector, this form of benchmarking might spur agencies to identify role models whose performance they would seek to emulate.

Performance Auditing

Performance auditing subjects statements by governments or agencies concerning their performance to review by auditors. The Chief Financial Officers Act of 1990 gives impetus to this concept by requiring agencies to include data on program performance in the supplement to their annual financial statements. The role of auditors in reviewing performance is not yet well defined, but if financial auditing is a guide, statements on performance will be a management responsibility. The task of auditors will be to review these statements for reliability and accuracy.

Performance-Based Budgeting

PBB links resources to outputs promised (or produced). The linkage can range from a lockstep relationship in which a unit of output is funded by a unit of resources, to one in which the budget merely identifies expected results associated with the requested or approved spending level. Governments that have formatted their budgets to show the units of outputs for the year have made the most progress in PBB. The State of Texas adopted a version of PBB in 1991, when its legislature removed standard line items from the appropriation bill and replaced them with performance targets. While many of these targets are process rather than output measures, the Legislative Budget Board is developing more advanced measures for the next biennial budget. In testimony before the Senate Governmental Affairs Committee, Frank Hodsoll, Deputy Director of OMB, suggested that the time is not yet ripe for PBB. "Performance measures must be agreed [to] and set before we attempt to set performance goals or standards. And performance goals and standards must be agreed [to] and set before we attempt performance budgeting." {Footnote 6} Many Federal agencies have reliable measures; the next step will be to establish standards, which might also be the appropriate next step for A.I.D.

Conclusion

U.S. Governments have taken many roads to performance management. In the past, many settled for performance measures. Today, however, avant garde governments are making bolder use of performance indicators to inform citizens on results, chart future courses of action, review what has been accomplished, and distribute resources. Somewhere within this spectrum, A.I.D. should be able to devise workable practices that expand Agency capacity without promising more than can be delivered. Although merely measuring performance will not be enough, A.I.D. already produces many measures tightly linking results and resources might go too far in removing critical judgment from country and project decisions. A key factor in determining the suitable niche for performance measurement is A.I.D.'s experience with similar systems.

3. Allocating A.I.D. Resources

For more than 30 years, A.I.D. has lived an examined life. The Agency has been repeatedly studied, reviewed, evaluated, and changed. Introspection has been customary inside the Agency, and outside observers have conducted countless studies. Some studies focused solely on A.I.D., others examined the Agency as part of broader reviews of U.S. foreign policy. Government entities have sponsored many studies, but more have been undertaken by interest groups, researchers, and others with ideas on how U.S. assistance can be made more effective. Within the past decade, A.I.D. has been put under the microscope by the Carlucci Commission (1983) and the Hamilton-Gilman Task Force (1989), as well as by its own Blue-Print for Development (1985), Development and the National Interest (1989), and reorganization and other management initiatives (1991). In 1992 alone, a Presidential commission recommended that A.I.D. be merged into the State Department, an OMB SWAT team scrutinized the Agency's operations, and the Government Accounting Office (GAO) conducted a sweeping management review. In addition, the GAO and the Inspector General have buffeted A.I.D. with a steady flow of critical reports, many of which have been accompanied by recommended changes in the Agency's operating procedures. {Footnote 7}

Every country in which A.I.D. operates--there are now more than 90--has been a venue for analysis. Countless projects (almost 2,000 are currently underway) have been evaluated to determine what has been accomplished and what lessons can be applied elsewhere. A.I.D. has been invented, reinvented, and invented again. Foreign aid emphasized technical assistance in the 1950s, infrastructure and capital investment in the 1960s, basic human needs in the 1970s, and open markets and institutional reform in the 1980s. These shifts have been accompanied by an incremental increase in the objectives A.I.D. has been directed to serve, to where it now is supposed to pursue almost 40 objectives (Report to the President 1992, 52; House Committee on Foreign Affairs 1989). Typically, the added objectives have not been accompanied by additional resources, but they have engendered persistent

questioning of the results of U.S. assistance.

Whatever ails A.I.D., it is not lack of attention. But judging from the harsh criticism the Agency continually receives, one could conclude that A.I.D. has not been adequately studied and that one more study would make a difference. It is not so. There is no basis for expecting yet another study or warmed-over recommendations to turn the tide. Arguably, the numerous studies and proposals contribute to the demoralization of A.I.D. staff and distract it from the difficult work of development.

Why does an Agency that has been studied so often appear so troubled and in need of fundamental restructuring? The answer lies largely beyond the scope of this report and can be traced to other problems, some of which are self-induced, such as inadequate internal controls that cannot cope with the Agency's far-flung operations; American dislike for foreign aid, especially since the Vietnam War experience and reinforced by fiscal stress; congressional micromanagement that has constrained the Agency; the slow, uncertain pace of development and the fact that few of the countries added to the assistance roster since 1961, when A.I.D. was established, have graduated to economic independence; an embedded management culture that is highly decentralized and difficult to change; and resource levels that have not kept pace with demands on the Agency.

Some of the conditions giving rise to these problems are clearly beyond A.I.D.'s control, but this does not mean the problems should be disregarded. It would be futile to upgrade A.I.D.'s technical capacity to allocate resources on the basis of performance unless its operating culture, including relations with Congress, is radically changed. Part of the transformation will require that A.I.D. confront certain matters within its purview, for example: (1) its programming process for determining the amount to be invested in each country and the projects to undertake, (2) its evaluation of completed and ongoing projects, and (3) its process for establishing performance targets and monitoring progress in achieving them. Each of these processes has a bearing on A.I.D.'s allocation of resources, but they have generally been divorced from one another. While all three have been part of A.I.D.'s management toolkit since the Agency's inception, each has had a distinct chronology in gaining prominence. Programming emerged first and has been the Missions' principal means of conducting policy dialogues with A.I.D./Washington and recipient countries. Evaluation was not strongly emphasized at the beginning, but gained prominence during the 1970s and 1980s. Measuring performance has always been stressed, but has recently become the organizing principle for management change in the 1990s.

Country Programming

Programming is how Missions and A.I.D./Washington establish an informational and analytical framework for making development policy. Programming is organized around field-assembled documents analyzing each country's situation and relating the situation to

U.S. or A.I.D. interests and objectives. Programming typically focuses on medium-term prospects (3 to 7 years) but also looks at decisions for the next year or two. In the field, programming is designed to stimulate discussion between Mission staff and host country counterparts. In Washington, programming culminates in a week of intensive review of medium-term activities and allocations.

The programming process has been transformed repeatedly over the years, and it has been periodically relabelled to reflect new priorities or concerns. But two features have endured throughout A.I.D.'s history: One is the country as the 'program package' or unit of analysis, and the other is the decentralized, mostly bottom-up, process for assembling program documents and proposals. These features have enabled A.I.D. to take advantage of field-based expertise, but they have greatly inhibited the capacity of headquarters to shape the programming process.

Programming in Historical Perspective

When A.I.D. was established in 1961, it inherited an elaborate, highly developed programming process from its predecessor, the Economic Cooperation Administration. The process was geared to produce a detailed program book, with an immense amount of economic and social data about each assisted country. Every major overseas Mission had its own program staff to undertake sectoral analyses, plan various types of assistance and development projects, and assemble the material into documents called Country Assistance Programs (CAPs). In the early 1960s, A.I.D. "had more skills and sophistication in programming and analysis than any other agency in foreign affairs" (Mosher and Harr 1970, 47). In fact, when the State Department launched its own programming process in 1963, the effort was led by former A.I.D. officials.

An important feature in A.I.D. programming was the linkage of programming to annual budget decisions. This linkage enabled the Agency to adapt CAPs to the governmentwide PPB system introduced in 1965. In line with PPB specifications, A.I.D. added some new documentation, such as program memoranda, which discussed overall policy objectives and how specific programs fit into them. When PPB withered away in the early 1970s, the program memoranda were discarded, but A.I.D.'s programming system remained intact. Programming at A.I.D. also survived the transition from a Democratic to a Republican White House. A transition task force (the Perkins Committee) recommended that A.I.D. be replaced by a new agency, but it expressly urged that country programming be retained as the means by which the agency 'weights priorities for development goals, examines alternative approaches to attaining them, defines needed self-help measures, and estimates requirements for external assistance' (President's Advisory Committee 1968, 17).

Spurred by the downsizing of A.I.D. in the early 1970s (staffing levels were cut in half) and the shift in emphasis from big infrastructure projects to basic human needs, the Agency adjusted its programming process in 1972. AIDTO Circular A-543 provided for a new document, a Development Assistance Program (DAP), for

each assisted country. DAP was to identify priority development problems, provide sectoral analysis of the problems, and recommend a multiyear assistance program. Although DAP purportedly was to overhaul the programming process, it was based on entrenched practices, especially country-centered, field-initiated studies and proposals. Although it kept many of the old objectives, DAP, like other A.I.D. reforms, was introduced to invigorate the programming process and prevent it from becoming a paper exercise, to winnow out some of the procedures and documentation that had built up over the years as requirements were added to the process, to infuse the process with fresh insights and perspectives, and to ensure that the latest priorities and aims were emphasized.

After the 1977 Presidential transition to the Carter Administration, A.I.D. once again revised its programming process, this time to emphasize the "new directions" assistance to the poor. An Agency task force criticized the existing bottom-up process on the grounds that it perpetuated previous allocation patterns and impeded A.I.D.'s response to current objectives (Program Procedures Task Force 1977). A.I.D. introduced an Indicative Planning Allocation (IPA) early in the programming cycle. The IPA, which was set by A.I.D./Washington, determined a provisional allocation of funds to each assisted country for 5 years. The allocation was to be based on quantitative measures of need and of the host country's commitment to address the underlying causes of poverty. With the IPA as a starting point, each Mission prepared a new programming document--a Country Development Strategy Statement (CDSS)--that was to be updated annually. The main purpose of CDSS was "to express the Mission's objectives, policies and programs A.I.D. should pursue, and explain the reasoning behind the choice" (A.I.D. 1978). CDSS enabled field Missions to plan for the next 5 years and to argue for higher allocations if they thought the IPA was inadequate. The early CDSS documents generally proved disappointing. Many Missions were unable to measure the extent of poverty or the progress made in mitigating this scourge. By the early 1980s, CDSS was modified to place greater reliance on qualitative indicators of economic and political conditions. {Footnote 8} This change reflected the difficulty of expressing results in measurable terms, as well as A.I.D.'s growing interest in linking assistance to economic and political indicators.

A.I.D. also modified the IPA to base it on the previous year's CDSS. This adjustment eased tensions between headquarters and Missions and reinforced bottom-up control of the programming process.

CDSS was periodically revised during the 1980s, but its country-based, decentralized structure remained intact. In fact, under the delegation of authority doctrine emphasized during the decade, Missions gained more authority to approve projects, while A.I.D./Washington's capacity to analyze policies and programs was weakened. With the spread of earmarks and the growing prominence of security assistance in lieu of development funds, A.I.D. was impelled to emphasize country allocations rather than program

results. Moreover, it appears that the increased use of economic and political indicators to gauge a country's openness to development downgraded the programming process as the means for allocating assistance. By the end of the 1980s, CDSS was more a reference document than a basis for action. It was a useful compilation of data, assessments, and other material on a country's political and economic development, but it did not offer much guidance on the amount or types of assistance that should be forthcoming. As the utility of CDSS diminished, A.I.D. placed greater reliance on Action Plans that linked short-term funding decisions to programs and objectives.

Evaluating Projects

Textbook models of programming systems typically show the evaluation of past actions feeding into decisions about the future. Although such models have been A.I.D.'s ideal for the past 30 years, implementing them has been difficult. As in many other organizations, programming and evaluation in A.I.D. often proceed on separate tracks, with assessment of results institutionally divorced from program and budget decisions.

A.I.D. sought to avoid this fate during its formative years by integrating responsibility for evaluation with other management functions. The Presidential task force that proposed the establishment of A.I.D. in 1961 grappled with the question of whether evaluation should be entrusted to a special unit or assigned to various program staffs. The task force subscribed to the view that "it is essential that the Administrator have the capability to assess periodically the substantive quality of progress made and results achieved under the various programs of the Agency" (A.I.D. 1961, 165). But task force members disagreed on how to institutionalize this capability. Some argued that evaluation should be consolidated in a separate office that would "make one official primarily responsible for all facets of program, personnel, and implementation evaluation" (A.I.D. 1961, 166). The prevailing view, however, was that program offices and other A.I.D. units share responsibility for undertaking evaluations and applying findings. The architects of A.I.D. expected that field Missions would conduct evaluations as part of their review of completed projects and design of new ones, that regional bureaus would do so in the course of developing action programs and allocating resources, and that headquarters staffs would perform evaluation as part of their responsibility for charting A.I.D.'s overall policy and direction. With nearly all A.I.D. units involved in evaluation, it was believed that this critical function would be amply attended to.

Whatever the virtues of this decentralized arrangement, evaluation was neglected in A.I.D. With everyone responsible for evaluating projects and programs, no one was in charge to argue for using the findings when plans were laid and funds allocated. A.I.D. faced a predicament that has beset many other government agencies: if evaluation is centralized, everyone else can ignore it; if evaluation is assigned to program units, these units tend to be more interested in looking at opportunities for the future

than at what was accomplished. When A.I.D. dispersed evaluation responsibility, a wide gap opened between doing and using evaluation. Although field officers dutifully prepared detailed project assessments, their work was often ignored during A.I.D.'s programming cycle.

The shortcomings of A.I.D.'s evaluation work were noted in a 1965 study of Agency operations conducted by the consulting firm of Booz-Allen & Hamilton:

It has often been said that AID has an inadequate memory. Evaluation reports on projects in process or completed are scarce...In spite of substantial evaluation efforts, frequently on an ad hoc basis, it is still true that AID has not yet developed a systematic process to appraise the consequences and results of its program operations and to exploit the rich accumulated experience of the agency. (p. 38)

The Booz-Allen team concluded that the main problem with A.I.D. evaluation work was "the absence of a clear focus of responsibility for coordinating such efforts and assuring effective follow-up and use of findings" (Booz-Allen & Hamilton 1965, 13). But the team rejected the notion that there should be a single focus for conducting all evaluation efforts of the Agency, saying "there are many good and sufficient reasons for many evaluation efforts by separate A.I.D. offices." To provide a comprehensive focus for evaluation while encouraging field and headquarters staffs to conduct their own assessments, Booz-Allen recommended that "[Bureau for Program and Policy Coordination] be designated as the office responsible for overview and coordination of this function for the entire agency" (Booz-Allen & Hamilton 1965, 100).

Despite frequent fine tuning and occasional reorganization, the Booz-Allen model has survived for many years. Overall responsibility for evaluation is lodged in the Center for Development Information and Evaluation (CDIE), which is now part of the Policy Directorate (formerly the Bureau for Program and Policy Coordination). But the bureaus have retained significant evaluation responsibility and each bureau periodically issues guidance to Missions in its jurisdiction. For example, the Africa Bureau, which has taken considerable initiative in devising new evaluation methods, has issued detailed guidelines for monitoring, assessing, and reporting on progress by Missions (A.I.D./Africa Bureau 1991; GAO 1991). Moreover, all Missions have ongoing evaluation responsibilities centered around the completion and review of projects.

This tripartite structure-CDIE, bureaus, and Missions-suits A.I.D.'s decentralized division of responsibility, but it encounters many of the management problems that have hobbled the Agency for years, especially the uneven attention to evaluation in Missions and bureaus, weak headquarters leadership and accountability, and a chronic failure to rely on evaluative findings in programming A.I.D. resources. {Footnote 9}

Geographically separated from most of A.I.D./Washington and

meagerly staffed, CDIE lacks the resources and clout to ensure that bureaus and Missions use evaluative findings. In November 1990, A.I.D. moved to strengthen evaluation by tripling CDIE's staff and budget. But the deficiencies in evaluation cannot be remedied by tinkering with this process alone. They are embedded in the operations and culture of the Agency. A.I.D. gets the evaluation that its organizational style and ethic dictate.

A.I.D. evaluations have continually been out of alignment with its programming process. The basic problem is that evaluation focuses on particular projects, whereas programming uses countries as its unit of analysis. The volume of project evaluations emanating from A.I.D. has been truly prodigious. In addition to the standard project completion reports, A.I.D. has issued a stream of reports drawn from project evaluations and other studies: for example, development experience abstracts, evaluation highlights, evaluation newsletters, evaluation occasional papers, impact evaluation series, program evaluation discussion papers, program design and evaluation methodology reports, and evaluation special studies.

Despite this flow of reports, evaluation has been disconnected from program and funding decisions. The 1983 Carlucci Commission, which reviewed U.S. assistance programs, noted that evaluations "are almost exclusively of technical nature...What is lacking is a comprehensive evaluation process that assesses the political impacts, interrelates the projects and programs...and expresses a judgment on their effectiveness in supporting U.S. interests" (p. 50). The Commission urged that evaluation criteria be incorporated into the development and selection of projects "because it is at this stage that program goals and the means for measuring their attainment must be established" (Carlucci Commission 1983).

A.I.D. has sought to broaden the scope of review from projects to programs by undertaking in-depth impact evaluations and, more recently, program assessments. Initiated in 1979 by CDIE, the impact studies have concentrated on the long-term effects of development assistance, especially the impacts on targeted beneficiaries, such as the very poor (White 1986). These studies look for changes that have ensued from development efforts in selected areas, such as family planning, and seek to determine whether U.S. assistance has made any difference. In contrast to project evaluations, which are typically managed by the Missions, impact studies have been carried out by CDIE-led teams spending 3 to 4 weeks in each of six to eight countries gathering data, observing results, and consulting with A.I.D. and host country experts. Findings from the impact studies have been disseminated within A.I.D. and to the larger development community through several publication series.

In 1991, CDIE initiated a program assessment series, as part of the A.I.D. Administrator's Evaluation Agenda. The program assessments closely examine the success of A.I.D. programs in key areas. The assessments are divided into program and policy evaluations and operations and management systems evaluations. The former includes topics in five areas: (1) economic policy

reform and public sector management; (2) private sector development; (3) agriculture, natural resources, and the environment; (4) human services; and (5) other programs. The latter includes topics on A.I.D.'s internal management and approach to development assistance.

Although A.I.D. has tried to apply the lessons of individual projects to a country's overall development, the task has not been easy. It has been complicated by the sheer number of projects'thousands are at various stages of development, implementation, or review at any time'but also by the shift in emphasis from infrastructure investment in the 1960s, to direct assistance to the poor in the 1970s, and to institution building and policy reform in the 1980s. This evolution to "softer" projects has made it more difficult to determine what has been accomplished. It is much easier to assess the impact of a dam or fertilizer factory built with U.S. assistance than to gauge what has been achieved by promoting the startup of community groups. Project evaluators can readily measure the volume of water diverted or fertilizer produced, or the number of acres irrigated or treated. They should be able to estimate the impact of these investments on food production and possibly on nutrition. But it is hard to assess the results of each of the thousands of small-scale activities funded by A.I.D. over the past decade or two. When development efforts seek improvement through institutional change, the lead time from assistance to measurable results is greatly increased and the assessment of benefits becomes more problematic.

Over the past two decades, the development community has become increasingly aware that improving a country's standard of living is a difficult, reversible process. Even when individual projects are successful, the assisted country often stagnates or regresses. Evaluation is most fruitful when there is a sense in the developing country of what works and why and when there is an understanding of how a particular activity contributes to larger objectives. This condition does not now prevail in much of the developing world. When countries that have received billions of dollars of donor aid are worse off than they were a decade or two ago, it is hard to evaluate specific projects in the light of a country's overall progress. It is now widely accepted that A.I.D. assistance is just one of the variables, and not always the most prominent, shaping the well-being of countries.

Current A.I.D. doctrine emphasizes economic and political institutions and policies as determinants of a country's capacity to effectively use external assistance. This emphasis is based on the expectation that countries with competitive markets and democratic institutions will progress more robustly than will countries with controlled markets and autocratic regimes. When institutional criteria become the basis for parcelling out assistance, the focus of evaluation inevitably shifts from the results of development to the preconditions for development.

Performance Management: Linking Programming and Evaluation

Programming and evaluation are A.I.D.'s principal tools for establishing objectives, allocating resources, and assessing results. Programming has always been a strong, but not necessarily effective, process in A.I.D. Evaluation has chronically been weak, even since the establishment of a central evaluation office. Programming needs periodic renewal, or it degenerates into a descriptive exercise of little value to policymakers. Evaluation needs periodic reemphasis, or its findings tend to be ignored. Both processes need to be more closely linked for programming to gain feedback from results to action and for evaluation to focus assessments on the objectives for which resources were programmed.

Two critical steps are necessary for linking programming and evaluation. One is to reorient programming from resources to results, the other is to reorient evaluation from a focus on individual projects to a focus on development results achieved at the country level. A.I.D. is moving in both directions by promoting a new performance management style that includes such initiatives as PRISM, a restructured programming process, PBB, and impact assessments. These initiatives are grounded in several fundamental assumptions. One is that objectives can be arrayed in hierarchical order, with strategic objectives defined by top management feeding into bureau objectives, which, in turn, guide the formulation of operational targets for field Missions. The hierarchy, variously called a "program logframe" or an "objective tree" establishes a logical connection between individual projects, activities, and a country's progress in meeting the objectives set for it.

A second, closely related assumption, is that objectives and results should be measured; hence the strong emphasis on performance indicators. There is agreement that performance measures should be applied in both programming and evaluation, but uncertainty lies in the extent to which performance measures should be used in allocating resources. Third, evaluation should focus on the results targeted when programs are decided. The expected results should be specified with sufficient precision to enable evaluators to determine whether they have been achieved. The final, and perhaps most important, assumption is that performance management should comport with A.I.D.'s decentralized mode of operation. Top-down guidance should benefit from bottom-up data and ideas. This premise was underscored in a 1991 A.I.D. study which insisted that "performance management must be constructed around the concept of country programming and A.I.D.'s decentralized operational structure." The study concluded that:

Performance management cannot be accomplished through "top-down" approaches, but requires productive collaboration with operational units through negotiating mechanisms that establish mutually agreed objectives and performance expectations in which operational managers have a clear commitment and stake.

A decentralized approach to performance management takes advantage of the progress already made in some units,

particularly the Africa Bureau, but runs the risk of being undermined by indifference at operating levels as well as by uneven implementation. It holds top managers hostage to cooperation and information at lower levels and can therefore become just another thwarted innovation rather than a blueprint for behavioral change. The architects of performance management hope to avoid these pitfalls by strengthening senior managers' capacity to strategically guide the Agency, relying on quantitative measures to program resources and monitor performance and encouraging Missions to cooperate by giving them greater operational flexibility.

Performance Management

For the Agency to make A.I.D.'s vision a reality, its organizational culture must be transformed "from one which focuses on design and obligation toward one which emphasizes and rewards implementation and development results" (A.I.D. 1991). This transformation is summed up in the concept of performance management. A.I.D. envisions managing for results at all levels of the Agency. Country-level decisions will provide the opportunity for senior managers to establish strategic objectives, taking account of each country's economic and social trends and its climate for development. Program decisions will emphasize projects and other activities that serve approved objectives. At operating levels, performance data will spur improvements in efficiency, internal control, and other management practices.

Performance management is an amalgam of new (and newly used) data along with changes in behavior. Obtaining data will be the easy part, getting managers to act on the basis of performance will be much harder, will take longer to effect, and may not be sustainable if interest or support abate. It will be some time before a fair assessment can be made of A.I.D.'s performance innovations. For now it is appropriate to note the changes being made through the three main performance management initiatives: the programming system, PRISM, and PBB.

The Programming System

In 1991, A.I.D. reviewed its programming system as part of its performance management study. The study argues for decentralized programming, but calls for a restructured system in which resources are invested and results evaluated in terms of measurable impacts on assisted countries (A.I.D. 1991). The restructured process is to be based on a division of labor that will give the Policy Directorate the capacity to define long-term objectives, enable the regional bureaus and their Missions to establish benchmarks against which performance will be measured, and provide medium-term targets consistent with agreed-on objectives. The programming system will also have procedures for review of country activities to ensure conformity to approved objectives, along with procedures for monitoring and reporting on performance.

The restructured programming system is modeled on initiatives taken by some regional bureaus. Taking advantage of the flexibility accorded the Development Fund for Africa, which is exempt from earmarks and other constraints imposed on other A.I.D. programs, the Africa Bureau has directed its major Missions to prepare an annual Assessment of Program Impact (API). API links specific projects to country strategies through a program logframe that specifies the performance targets for each year covered by the program. In targeting future results, Missions specify baseline performance levels as a starting point against which progress is measured. With the baseline and targets in place, API showcases one of the strengths of decentralization: a bureau (or Mission) can take initiative without waiting for the rest of the Agency to act. There are disadvantages, however, such as uneven progress and a weak capacity of top leaders to implement needed changes.

The recent revisions in the programming process are neither the first nor the last; nor are they the first to call attention to results. Like previous reforms, the recent revisions will make a difference only if they are accompanied by behavioral changes.

Program Performance Information System for Strategic Management

The programming reforms depend on the availability and use of performance information at all echelons of the Agency. PRISM is A.I.D.'s initiative for compiling and classifying performance measures that inform managers on whether expected results are being achieved and country program outcomes are moving in the intended direction.

PRISM is not being built from scratch. It is premised on the notion that bureaus and Missions have considerable performance data and that it is appropriate to base an Agencywide system on preexisting capabilities. PRISM culls through hundreds of objectives and thousands of indicators identified in country programming documents and organizes them into a few clusters²⁰ at the outset, perhaps half as many in the final system. Once selected, the clusters are used for classifying Mission and bureau objectives into Agencywide categories that permit cross-country comparisons, determinations about the programs overall success in meeting development goals, and evaluations of programs and their impacts and cost-effectiveness.

PRISM is designed to impose some top-down logic on what will continue to be largely a bottom-up arrangement. As PRISM matures, it should generate some new data needs and is likely to exert some pressure on Missions to align their objectives with bureau or Agency objectives. PRISM should also enhance the capacity of central units to review Mission objectives and activities. Having information does not ensure that an organization will be transformed into one that embraces a performance ethic. Unless information is used for such critical purposes as establishing programs and allocating resources, nothing will change. For this reason, A.I.D. is exploring the possibility of implementing a PBB system.

Performance-Based Budgeting

Like other initiatives in A.I.D., PBB began in a few bureaus and is now spreading throughout the Agency. The Latin America and Caribbean (LAC) and Africa Bureaus have pioneered PBB systems of their own, and A.I.D./Washington is exploring the feasibility of Agencywide implementation.

A.I.D. has approached PBB more as a means of assessing the probable effectiveness of assistance than as a means of relating resources to results. Although the PBB systems in the LAC and Africa Bureaus differ in the weights assigned critical variables, both emphasize the economic and political climate for investment rather than the effectiveness of A.I.D. assistance. The PBB system in the Africa Bureau also measures the need for assistance, but neither Africa nor LAC tie level of aid to program outcomes (Development Alternatives 1991).

The LAC Bureau uses four sets of indicators pertaining to economic, social, environmental, and political conditions. The various conditions are assigned quantitative weights (totaling 100 percent) on the basis of which the Bureau judges the level of assistance appropriate for each country. Economic performance accounts for 50 percent of total weight and consists of six indicators: macroeconomic policy, financing of the deficit, exchange rates, trade policy, the regulatory environment, and price controls. Environmental policy accounts for 10 percent of a country's rating and is drawn from a questionnaire on environmental management and institutions. Political condition (20 percent) focuses on human rights and commitment to democracy. Only social performance (20 percent) uses program measures. These pertain to school enrollment, infant vaccination rates, and contraceptive practices.

The Africa Bureau uses a system that combines measures of performance, need, and country size. Performance is based on macroeconomic policies (25 percent of the total score), on 'equity' as reflected in agricultural, health, education, and family planning policies (12.5 percent), and on political policy (12.5 percent). The remaining 50 percent of total weight is allocated on the basis of country size and need, as measured by child and infant mortality.

The A.I.D. Policy Directorate has developed a resource allocation model that builds on work in the bureaus and emphasizes macroeconomic policies. Once installed, the new system will be used by all bureaus, thus imposing a degree of uniformity unusual for A.I.D.

Conclusion

The several performance management approaches applied in A.I.D. do not conform to what PBB is commonly understood to mean. In contrast to conventional PBB systems that focus on planned or budgeted results, A.I.D.'s systems measure conditions for

investment, sectoral policies, or program inputs. The A.I.D. approach is predicated on the expectation that the success or failure of assistance hinges on the policy environment within which development is taking place. Whatever the validity of this argument, it is not PBB. It incorrectly labels inputs as outputs and mistakenly treats the conditions for development as the results of development. Of even greater concern is the fact that performance in the PBB system is not consistent with performance in PRISM or in A.I.D. programming systems, which view performance as specific outcomes or impacts that can be linked directly to A.I.D. investments. The investment formulas devised in A.I.D. may provide a basis for determining the level of assistance in a country. They do not offer any guidance on how these resources should be programmed nor on the particular sectors or projects to be assisted. A genuine PBB system must connect resources used to results achieved. How A.I.D. might go about devising such a system is the subject of the next section.

4. Designing a Performance-Based Management System for A.I.D.

The language and promise of performance pervade A.I.D. At all levels, exhortations about managing for results, establishing objectives and measuring progress, monitoring impacts and outcomes, evaluating the difference that U.S. assistance makes, and so on abound. Performance is promoted in Congressional Presentation documents, Administrator's memoranda and directives, regional bureau guidance, programming documents, and Mission and project reports.

Exhortation, although needed, is not enough, however. It must be supported by two critical changes in Agency behavior: making A.I.D. more inclined and better able to take performance seriously and making results--actual and projected--the basis of program and resource decisions. The first requires truly far-reaching changes in Agency operations and practices to remove obstacles that inhibit results. The second entails implementing some features of PBB. The two reforms are interdependent, for without a transformation of A.I.D.'s operating culture, efforts to install PBB will be futile. And without procedures for using performance data and assessments in making important decisions, A.I.D. will inevitably revert to its old ways.

Many of the arguments presented in this section are not new; they have been voiced repeatedly in government reports and other assessments. The fact that the arguments have not been adopted suggests the difficult task facing A.I.D., both with respect to changes requiring congressional acquiescence and those dependent on A.I.D.'s managerial initiative. But the arguments warrant restatement, because without far-reaching reforms, performance management will be a hollow promise and the stream of adverse reports will continue to flow.

Recommendations for a Performance-Based Budgeting System

PBB is viable only in an agency that manages for results and demonstrates its seriousness about performance by specifying

objectives with sufficient precision for informed programming and resource decisions and by routinely monitoring outcomes and taking corrective action when outcomes veer from course. The performance management and measurement systems being introduced in A.I.D. are preconditions for PBB. The discussion that follows assumes that these systems are already in place. It therefore concentrates on the elements of PBB.

PBB should be used principally to allocate funds within-not between-countries. A.I.D. makes two basic resource decisions: how to allocate funds to approximately 90 countries and, within each assisted country, how to dispense funds to approved projects. The PBB approaches devised by the LAC and Africa Bureaus, as well as the Agencywide system introduced by the Policy Directorate, are oriented to allocation decisions. These approaches rank countries on the basis of political and economic criteria, with higher ranked countries receiving more assistance. Although a statistical formula for parcelling funds among countries has obvious appeal (it purports to take politics and subjective judgment out of the process) and has been sought by Congress, it is a misguided application of the PBB concept. It would be more productive to use performance-based criteria for allocating funds within countries and to rely on informed judgment and measures of "investment climate" for dividing funds among countries. This conclusion is based on four considerations.

1. The factors involved in current formulas for allocating funds address each country's readiness (and in the case of the Africa Bureau, the country's need) for development, not what is accomplished with U.S. assistance. Arguably, performance should always denote the results ensuing from governmental action, not what might happen in the absence of such action. Furthermore, performance-based allocation should distribute resources based on actual or projected outcomes, not on conditions prevailing before assistance is provided. In fact, A.I.D. formulas allow countries with lower performance (e.g., less progress in reducing infant mortality) to score higher than countries showing superior performance. This perverse ranking occurs when a country with favorable economic or political preconditions of development fails to effectively implement the projects for which assistance is provided.

2. In any formula-based allocation, small changes in the weights assigned various criteria can make a big difference in country rankings. Attributing 40 percent of the total score to macroeconomic policies will yield a different ranking than if 50 percent were assigned to this category. Rankings also change if they are adjusted for country size, measures of need, and other factors. It is wrong to impute objectivity to scores that are based on subjective judgments of the relative value, criteria, or weights included in the formula. Countries often are bunched together in the rankings, so that slight adjustments in the formula can greatly enhance or retard a country's fundability. For example, 37 of the 90 countries the Policy Directorate ranked on the basis of economic performance indicators fall within a 15-point range on a 100-point scale. When countries are clustered

so closely, it may be inappropriate to use the rankings to allocate assistance.

3. It would be imprudent (and probably impractical as well) to ignore political assessments and professional judgments in allocating assistance. There is nothing amiss in investing more in low-ranked countries in which the United States has vital interests than in higher ranked countries in which U.S. interests are less salient. In fact, consideration of U.S. interests is the basis for making Egypt the second biggest recipient of U.S. aid, even though it falls in the lower half of the country rankings.

4. During the interval from country programming decisions to project completion-typically longer than 3 years-the conditions on which the rankings are based are likely to change, sometimes dramatically. Some countries that ranked high when allocations were made will have much lower scores after funds have actually been used. In such cases, the performance budgeted at the time of resource allocations will have no relevance to the performance delivered after project's completion.

There is a useful niche in A.I.D. for formula-aided allocations, but this should be distinct from PBB. PBB should be a means of allocating resources to programs and projects within countries and should be based on actual or expected performance, not on preconditions of development.

Economic and political criteria should serve as aids to judgment in determining country allocation levels, not as rigid allocation formulas. Although PBB should be used for making program and project decisions, A.I.D. still needs a method for determining the amount of assistance for each country. The allocation of resources to a country is inherently, but not exclusively, a political decision. It entails judgments concerning U.S. interests, the effectiveness of previous assistance, the future course of the country's development and its capacity to absorb external aid, the stability of its regime, the durability of its commitment to open markets, and the roles and contributions of other donors. These judgments cannot be reduced to a single formula.

Yet, it would be inappropriate to disregard measurable qualities of development in setting levels of assistance. Although country allocations require political judgment, politics should not be the sole consideration. It would be foolhardy to ignore a country's climate for investment, as measured by economic policies and political conditions, or its need for external aid, as measured by the prevalence of poverty or other adversities. A.I.D. must consider the many relevant characteristics of development while allowing appropriate scope for political and professional judgments. This study proposes a three-step procedure for making country allocations: (1) A.I.D. officials should prepare measurements of each assisted country's development potential and development need; (2) they should then prepare a matrix, listing the assisted countries according to those measures; and (3) the Agency and bureaus should make

judgments about the amounts to be provided each country.

In step one, development potential and development need are each scored separately; at no time should the scores be combined. Development potential consists of economic criteria similar to those used by the Africa Bureau, the LAC Bureau, and the Policy Directorate, and some political measures. The weights assigned to the various criteria are a matter of judgment, but they should be uniform throughout the Agency. Moreover, a portion of the score should be reserved for Mission (or bureau) assessments of how effectively each country can use U.S. assistance. All countries should then be ranked on the basis of their development potential score.

A similar procedure should be used for constructing a development need index. Appropriate factors might include infant mortality, per capita income, and access to education and other social services. Here, too, a portion of the score should be reserved for professional judgment. All countries should be ranked on the basis of their development need score. A scale factor is also necessary to take into consideration the size of the country. The next step is to array the rankings in a two-dimensional matrix, along the lines suggested in Table 1. Each index should be divided into quintiles, with each quintile consisting of one-fifth of the assisted countries. Arraying the two indexes in a single matrix provides 25 combinations (or boxes). Each box lists only one country, as determined by the country's potential and need scores. The matrix thus provides a useful visual summary of the statistical data.

The final step is determining country allocations. Countries in the upper right portion of the matrix show the highest potential and need and therefore should presumably qualify for assistance. Many countries are likely to be clustered in the middle of the matrix. Although A.I.D. officials should give the matrix and the underlying scores careful attention in making their allocation decision, they are not bound to the distribution patterns indicated by the matrix. A.I.D. officials should be required, however, to justify allocations to countries placed in the lower left part of the matrix, those with low potential or need.

Country allocations should be finalized on the basis of program and project objectives and targets. Allocations to countries cannot be decided independently of program and project decisions. Regardless of how favorably a country measures in terms of development potential and need, A.I.D. must still determine whether its resources can be put to productive use. This determination should be made at two stages of Agency operations: during programming and in the course of project management. Every country allocation made on the basis of development potential and need should be regarded as tentative until the programming cycle has been completed. During this cycle, country objectives enter the picture, along with targets for measuring progress both essential to the PBB approach to allocating A.I.D. resources.

The Agency has three basic options for distributing resources on the basis of performance. The first is to structure an Agencywide competition in which countries budgeting the highest performance levels are favored, regardless of where they are located. The second is to structure competition within regions. And the third is to assess performance on a country-by-country basis, without any cross-country comparisons.

The Agencywide comparison of performance targets is likely to be unwieldy—it is hard to compare 90 countries on a single performance dimension—and would run afoul of congressional earmarks and the separate Development Fund for Africa. Regional comparisons, however, comport with A.I.D.'s bureau-centered programming system. But even within-region competition must come to grips with the difficulty of comparing countries at different stages of development. Consider two countries, one with high infant mortality and the other with a lower rate due to recent improvements in nutrition and health care. If funds are allocated on the basis of the prospective rate of improvement (or the rate of change), the lion's share will go to the first country and the second country will be penalized for its progress. If, however, funds are allocated on the basis of the absolute level of performance (such as the actual or projected mortality rate), the second country will be favored and the first country's compelling needs would be shortchanged. It is possible to combine the rate of change and level of performance into a weighted score, but this approach will only hide the problem of cross-country comparisons under statistical data.

The third approach—within country assessment—avoids this problem, but it does not satisfy A.I.D.'s interest in a PBB system that facilitates the allocation of funds among countries. In the view of this study, however, performance measurement and assessment must be done principally on a country-by-country basis and (as discussed in the next recommendation) should focus on what each project is contributing to development. Nevertheless, it is feasible to meld cross-country comparisons and within-country analyses through an interactive process in which the allocations made to countries pursuant to the economic and political criteria are tentative, and final allocations are based on each country's projected performance. This two-stage process should be concentrated in the regional bureaus, though it would be prudent for the Administrator to reserve some funds for Agencywide decisions and contingencies.

This study envisions a two-stage process, beginning with tentative allocations drawn from the previously discussed assessments of development potential and need. What makes this allocation tentative is that Missions will not be assured funding until they have specified the performance they promise to deliver with the requested funds. The competition in the second stage, and the dialogue between the Mission and A.I.D./Washington, should be riveted to the results that will be purchased with the funds. If the process works as envisioned, the tentative allocations will be modified to shift funds from lower to higher performing Missions and projects.

These final allocations, like the tentative ones drawn from measures of development potential and need, should be based on the informed judgment of A.I.D. officials, not on strict fidelity to statistical scores. PBB must be a means of improving program and budget decision, not a substitute for judgment.

The competition for performance-based allocations should focus on country programs and major projects. The process of finalizing country allocations requires a sharp focus on what projects are contributing to performance. This focus is not yet sufficiently emphasized in A.I.D.'s emerging program performance measurement system, which is keyed to country program performance. Attention to the pace and results of country development enables the Agency to consider both past and promised performance in planning and carrying out multiyear programs and objectives. A.I.D. must always keep its eye on the prize: improvement in the well-being of a country and its people and projects and other assistance (such as food aid) are A.I.D.'s means of delivering it. Currently, however, inadequate attention is being paid to the contributions made or expected from individual projects.

Without a logical connection between projects on the one hand and developmental results on the other, few of the many Agency expenditures will be justified. A.I.D. has sought to forge this connection through logframes linking specific project targets to broad program objectives. Current guidelines instruct Missions on the types of measures to be used and emphasize that project funding should be based on projected outputs and impacts. But these guidelines do not go far enough. Still lacking is sufficient follow through and commitment to demonstrate that A.I.D. takes performance seriously and is willing to program resources on that basis, even to the point of withdrawing funds from slow-producing projects.

The test of performance is always in results, not in plans. When an agency stresses planned performance but neglects actual results, it induces bidders for resources to inflate their expectations and to promise more than they can deliver. Such spenders will be rewarded with more funds and will not be penalized when results fall short of expectations. Only when an agency rigorously monitors and evaluates actual performance, compares results with expectations, and makes appropriate adjustments- including midproject changes in funding levels-are spenders likely to take performance seriously.

The trick, therefore, is to program and budget on the basis of expected project results and for bureau and Mission managers to closely monitor actual results. These officials will have data showing how different project mixes and funding levels will impact on projected performance. This information will facilitate trade-offs between similar projects in the same country, between different projects (such as nutrition and health care) in the same country, or between projects in different countries. Each type of trade-off requires a somewhat different mix of data. Comparing similar projects in the same country calls for data on

alternative funding levels and the results expected at each level. Ideally, the results should be expressed in absolute terms and as rates of change; both should be displayed for a full programming or project cycle, often from 3 to 7 years.

A more complicated trade-off occurs in comparing different projects in the same country. The procedure recommended here is to display spending and results associated with each type of project separately, without reducing them to a common objective or to estimated dollar benefits. For example, if health care and education are being compared, the projected results of each, at various funding levels, should be displayed, but the two sets of results should not be combined in a single measure. Although health care and education share some development objectives, such as reduction in infant mortality, they operate in different time frames and have some distinct objectives. Inasmuch as PBB is envisioned as an aid to judgment, it must express each sector's objectives in the sector's own terms. The measures should not be reduced to cost-benefit (or cost-effectiveness) ratios for intersector comparisons.

Similar advice pertains to comparisons between countries. It is far better to display each country's performance separately than to homogenize it in weighted scores that conceal the assumptions on which they are based. Only when the budgeted outputs of various countries serve identical ends should they be combined in a single score.

Performance should be measured in terms of change from baseline or current levels to projected or targeted levels. A measure of performance should always measure change from one period to another, from one level of funding to another, or from one set of government programs to another. To say that a program is performing is to conclude that it is producing outcomes that are different from those that would have prevailed in the absence of the program. This is the essence of change-oriented performance measurement.

PBB takes this concept further by allocating resources on the basis of the changes to be purchased with government funds. Change is rarely an all-or-nothing proposition; typically it entails marginal differences in results, depending on the resources allocated and the success of the government's intervention. It should be feasible, therefore, to measure the change associated with different funding levels and to determine whether the results anticipated when the budget was decided have been realized.

Although change is an inherent feature of performance, it is especially critical in development. Development is itself a process of change; it has no meaning except in the context of differences between a country's past and future condition. The concept of change is essential to A.I.D. because it focuses the Agency on its overriding purpose, which is to further improvement in the well-being of developing countries and their people.

Toward this end, PBB, as well as other relevant A.I.D. initiatives, should emphasize change from current or baseline levels to planned levels. Outcomes or impacts should be expressed as measure of change, either in absolute terms or as percentages. Performance-based budgeting should facilitate the allocation of resources on the basis of projected change in outputs and outcomes. Properly designed, PBB is a form of marginal analysis in which incremental changes in resources are associated with incremental changes in results. This linkage is advisory, supplementing not supplanting the judgment of A.I.D. officials. Table 2 suggests a simple format for displaying the relationship between resources and results. In this hypothetical case, each successive spending level represents an increase in resources for programs intended to reduce infant mortality in a developing country. The first baseline level projects the number of families to be assisted and the infant mortality rate expected if the project is not undertaken or if enhancements are not made in a previously authorized project. Outputs and outcomes are separately projected (the two should not be comingled) for each of several years into the future. The multiyear projection is useful because assistance typically takes several years to produce planned changes.

Each spending level after the baseline projects a rise in outputs and an improvement in outcomes. These projections can be computed as rates of change, thereby allowing budgetmakers to calculate the estimated improvement associated with each increment of expenditure. The number and spacing of the increments is a matter of convenience; too many unduly complicate the analysis; too few make it difficult to link resources and change.

Table 2 also displays resources and estimated results for a single country. A similar format can be used for two or more countries with similar projects and identical objectives (see the example in Table 3). Both tables should be backed up with detailed supporting data. Note that each country shares the objective of reducing infant mortality through improving the nutrition of poor families. At this stage of its development, PBB should not be used to compare budgets and results for programs or countries with different objectives.

A.I.D. should clearly inform Missions and other suppliers of performance data on how the measures will be used in programming and budgeting. The accurate generation, analysis, and use of performance data are contingent on trust between the suppliers and users of the measures. The surest way of impairing trust is keeping producers of data in the dark about how the measures will be used, by whom, and for what purposes. This problem is widespread, because organizations often decide on the information they want first and only afterward consider the uses for the data.

Every performance measurement system has an implied threat that the data will be used against those generating it. The threat is intensified when the data feed into a PBB system used in allocating resources. PBB instructs field operators and others to

compile and report data that may be used to cut their budgets. The geographical distance between field and headquarters and differences in perspectives can aggravate these concerns.

There is no perfect solution to this problem; if there were, PBB would be used much more widely than it is. A.I.D. can take several steps to ease the problem, however, beginning with making firm assurances that performance data will not be used as a rigid formula for allocating resources. Equally important, A.I.D. officials should work to erase the distinction between producers and users of data. Generators of the data must be encouraged to apply the data in their work. Ideally, Missions should be the prime producers and users of performance data. They should apply performance measures in planning programs, justifying their budgets, communicating with bureaus and other headquarter units, reviewing projects, and assessing results.

The Agency must also take steps to ensure that higher organizational units take into account the measures that Missions and others prepare. Over the years, the performance measurement initiatives of Federal agencies have been repeatedly undermined by disuse and by the failure of senior management to use the data they requested. In such situations, it does not take long for operating managers to sense that their reports are being filed and forgotten. Nor does it take long for them to adapt to this reality by turning performance measurement into meaningless boilerplate.

PBB should enrich the dialogue among organization levels in A.I.D., especially between Missions and bureaus. The two sides should focus their discussions on what is expected to be accomplished and the results of past decisions. There is no better way to reinforce the centrality of performance in the Agency than by demonstrating that the people in charge care about performance.

Caring can be taken too far, however, if it rigidifies into a formula supplanting judgment. When this occurs, dialogue stops and spending units maneuver to protect their resources while paying lip service to performance measurement. The architects of PBB should be guided by a simple rule: Don't ask those who spend money to load the gun that will be pointed at them.

A.I.D. should guard long-term development objectives against short-term vicissitudes and setbacks. With performance measurement's emphasis on results, A.I.D. officials are tempted to build a system that focuses more on recent or impending achievements than on long-term objectives. Even when objectives are set for the medium term (3 to 7 years ahead), officials tend to concentrate on quarterly or annual results, as evidenced in periodic LAC and Africa Bureau assessments of economic and political developments in assisted countries. The implication of these assessments is that the flow of U.S. dollars should be conditioned on current events, along the lines of the adjustment policies enforced by the International Monetary Fund.

Except when the shifts are truly dramatic, such as the overthrow of a country's democratically elected leaders or collapse of its currency, A.I.D. should avoid adjusting U.S. assistance to economic or political fluctuations. The course of development is almost always uneven, with progress interrupted by occasional setbacks. A.I.D. should seek to contain the impacts of these reversals by maintaining a constant focus on where it hopes the country will be in 5 or 10 years.

Although A.I.D. cannot, and should not, ignore what is happening now, it should assess current developments in the light of future objectives. At every turn, the Agency's principal concern should be whether the country is on track to achieving the objectives targeted for it. The following suggestions should encourage a longer time perspective:

Monthly performance reports should be discouraged, although operational details (such as obligations, disbursements, staffing, and similar matters) should be reported each month.

Frequent reporting requirements reduce the examination of performance to a rote exercise leading to the 'file and forget' behavior mentioned earlier.

Quarterly performance reports should concentrate on key output measures, such as the number of persons served. More detailed reports, including impact indicators, should be reported annually. The different reporting periods arise from the difference between outputs and impacts. Reports on outputs measure the volume of services provided by U.S. assistance.

Reports on impact measure the effects of these services on the well-being of the populations, which typically take years to be realized.

Performance reports should regularly assess the impacts of current results on long-term objectives. The quarterly reports should estimate whether the outputs targeted by each project are still on track. Annual reports should project future outcomes.

A.I.D. Missions should systematically monitor and report on performance, carefully analyze variances from expected performance, and take corrective steps when appropriate. Program and project evaluations should focus on whether budgeted objectives and promised improvements in the country's well-being have been achieved. PBB will be of little value if it is confined to the preparation of the budget and the initial approval of projects and allocation of resources. The key to performance must not be whether results are promised, but whether they have been delivered. Accordingly, PBB should not only be a means of preparing the budget, but also the basis on which the budget is implemented. This entails the systematic monitoring and assessment of results, periodic comparison of results against budgeted objectives and targets, analysis of variances between planned and actual performance, and adjustments in the budget, activities, or objectives as indicated by the variance analysis.

Expenditures, activities, and results rarely turn out as planned. The analysis of variance concentrates on significant deviations from planned levels. Although businesses routinely use variance analysis, the public sector rarely does. The failure of government agencies to apply this standard practice calls into question their commitment to performance. If A.I.D. is serious about managing for results, it should build variance reporting and analysis into its performance measurement system.

The variance control process should focus on three related elements of performance: the amount spent, the outputs produced, and the progress made in achieving outcomes. For each of these elements, results should be compared with planned levels, and significant variances (i.e., those above a preset threshold) should be reviewed to determine whether corrective action is appropriate. Such action may include adjustments in the budget or in approved activities, as well as revision in targeted outputs and outcomes. Budget adjustments may be up or down and could entail the reprogramming of funds, redirection or termination of projects, and redefinition of objectives.

Variance analysis works only when results are reported shortly after the close of the period to which they pertain. A.I.D. should schedule output reports to be available no later than 10 to 20 days after the end of each quarter. Annual performance reports, including impact data, should be provided on a schedule that enables A.I.D. bureaus to feed the results into their programming and budgeting cycle. If the reports are delayed, the Agency is not likely to use them as the basis for corrective action.

Program and project evaluations should be turned into forms of variance assessments. Their principal focus should be on whether the objectives set for the activity being reviewed have been met. They should systematically compare promised and actual results and seek to explain differences between the two. Some standardization should be introduced into the format for project evaluations, especially with regard to the achievement of planned objectives.

Making Performance Management a Reality in A.I.D.

The effective use of the proposed initiatives requires effort and commitment at all levels of the Agency and a change in its culture. Without the necessary effort and cultural change, the initiatives are likely to be rhetorical substitutes for action rather than prods to reform. The problem is not a lack of commitment to performance. A.I.D. managers care whether their programs are performing as intended, whether assistance is making a difference, and whether promised benefits are being realized.

But new measurement and monitoring systems, accompanied by verbal support from the top, cannot overcome embedded impediments to performance management. It is imperative, therefore, that A.I.D. act to remove those impediments, which will require a cultural

change in the Agency.

Removing Impediments to Performance.

Almost every phase of A.I.D. operations--the development process, resource programming and budgeting, congressional involvement, and headquarters-field relations--inhibits sustained attention to results. Consider first the overall climate in which A.I.D. operates. The pace of development is uncertain, slow, and halting and is typically measured in decades rather than in the compressed periods spanned by strategic planning and multiyear programming. How can A.I.D. assess results when backsliding is pervasive, when beleaguered governments retreat from economic reform or democratization, or when exogenous conditions over which governments have little control, such as swings in commodity prices or sudden capital flight, overwhelm their best efforts? Even when evaluators find favorable results, they must regard the good news as tentative and subject to reversal. Adding to the difficulty, performance depends more on what the assisted country does than on A.I.D. action. Although the Agency's reputation is on the line, results of its actions depend on the performance of others. Moreover, A.I.D. contributes only a small, declining share of development funds. U.S. assistance is often part of a multilateral package in which it is not always possible to separate the impact of the U.S. contribution from that of others.

These known characteristics of foreign assistance are among the factors that have moved A.I.D. to focus on institutional conditions of development rather than on the results of development. But this is a poor substitute for consideration of what U.S. assistance is achieving.

The focus on performance is further clouded by congressional action. Over the years, Congress has steadily expanded the list of objectives that A.I.D. must pursue. It has squeezed foreign assistance so that the Agency has fewer real (inflation-adjusted) dollars to distribute than in the past. Few in Congress seem to care that objectives sometimes clash or that resources are spread too thinly to enable A.I.D. to pursue objectives vigorously. The failure of the Hamilton-Gilman bill a few years ago dampened congressional interest in rationalizing A.I.D.'s objectives and priorities. The multiplication of objectives has spurred interested groups to lobby for their share of limited development funds; the results has been a progressive rise in congressional earmarking of A.I.D. funds and a concomitant shrinkage in the Agency's operating flexibility. To make matters worse, even before assistance is earmarked in appropriations bills and reports, A.I.D. must go to Capitol Hill with a detailed list of countries and projects. A.I.D. must adhere to these itemizations except when it is directed otherwise by Congress or is authorized to reprogram funds.

Congressional involvement inexorably crowds out serious attention to performance. Legislated objectives offer little guidance on the direction the Agency should go, and earmarks prescribe the

allocation of funds regardless of developments in the affected countries. Congress demands performance and is frustrated by failure to get it, but Congress does not see the connection between its own actions and A.I.D.'s record.

Yet, much of the damage to A.I.D.'s capacity to manage for performance is self-inflicted. The Agency maintains an overextended pipeline of projects in various stages of development. For many projects, initiation or completion is years away, if it occurs at all. Virtually all new funds are mortgaged to previously approved projects, and new activities must queue up for funding that might be years away. The mortgage causes programming to be out of synch with the budget cycle. The former plans for resources that will not be available for some time; the latter is locked into ongoing commitments that do not always reflect current priorities. This misalignment discourages bureaus and Missions from focusing on performance and thwarts efforts to allocate resources on the basis of results. The performance directives streaming from headquarters come too late to have much impact on projects already clogging the pipeline and too early for new projects that will take years to move from the drawing board to fruition. The misalignment persists in the field, where managers begin their tour by inheriting old projects programmed years earlier and end their tour by bequeathing their planned initiatives to the next shift. For those coming or going, the objectives set for the various projects tend to be stale, alien, someone else's. In this environment, Missions cope with changing demand without really changing what they do.

A.I.D. values its culture of decentralization, but this is not a culture that ensures performance. It is a culture that stabilizes the Agency and eases friction between headquarters and the field. It is a culture that gives central officials an inordinate involvement in the details of operations, while curbing their capacity to make and carry through on policy. This relationship does not foster accountability for results at the top or sensitivity to centrally established objectives at the bottom.

Orienting A.I.D. to Performance

The essential first task in managing for results is to remove obstacles that have accumulated over the years. This reorientation of A.I.D. should proceed on two tracks: in Congress and in the Agency. A new legislative relationship is a prerequisite for other changes. Without congressional action, the Agency will have neither license nor backbone to transform itself. Congress is the source not only of the requirements and restrictions it legislates but of A.I.D.'s operating culture as well.

Congress should clarify A.I.D.'S development objectives and specify criteria for assessing progress. The hodgepodge of 40 or so legislated objectives is anything but a road map for the Agency. It gives lip service to all objectives, priority to none, and ensures that A.I.D. will fall short of the mark, no matter how hard it tries.

Through PRISM, A.I.D. has sought to rationalize its priorities by boiling down hundreds of Mission activities to a handful of broad objectives that can be related to congressional directives. These objectives serve as the basic clusters within which specific performance measures are grouped. While this initiative is welcome, without corresponding legislative action, it will not be long before new goals are prescribed, old ones become muddled, and the Agency is once again overloaded with objectives that offer little guidance for planning programs and allocating resources. To discourage this tendency, Congress should restate A.I.D.'s purpose and priorities along the lines proposed by the Hamilton-Gilman task force. The process of formulating attainable objectives should move Congress to rethink the aims of foreign assistance, the organization and operations of A.I.D., and the role of Congress. This reexamination should lead to legislation removing some of the rigidities that hobble the Agency.

It is not enough for Congress to prescribe objectives; Congress must also provide guidance for assessing progress. To do so will require Congress to state objectives with sufficient precision to serve as benchmarks for measuring results. It is not necessary or desirable for Congress to specify operational targets, but it should set the strategic measures to be used in assessing overall performance. If, for example, the goal is reduction in infant mortality, let Congress say so by tying this objective to authorizing legislation and annual appropriations.

The Congressional Presentation should focus on programs and performance and should not detail country and project funding levels. The Congressional Presentation is an old practice that enables Congress to closely monitor A.I.D. operations and to dictate spending priorities. Congress wants to know where public funds are going, which countries are being aided, and what projects are being funded. The offshoot is a bulging document that is long on spending details but says little about progress in achieving targeted benefits. The 1993 edition has a 780-page country list plus another 650 pages of project detail.

There is a fundamental contradiction between these detailed spending lists and genuine consideration of performance. When a country knows in advance that it is budgeted for a certain amount of assistance, negotiations with A.I.D. are likely to concentrate on operational details, not on promised results. Of what avail is it to discuss performance when Missions and host countries know how much will be available to them regardless of whether funded activities are living up to expectations? It is foolhardy to pretend that the Agency can manage for results when advance decisions are so detailed and unbending.

Country-Agency negotiations are at the heart of the process of orienting A.I.D. to performance. The Agency cannot produce by itself; it can get results only by acting in concert with the host country. A key aim of negotiations should be to clarify what the country will do to facilitate development and the expected impacts of U.S. assistance. It is especially important for

negotiators to take a close look at projects in the pipeline to be realistic about what can be accomplished and the rate at which development funds can be effectively used. These discussions already take place, but the negotiators may merely be going through the motions when they are locked into preset funding commitments.

The Congressional Presentation should be overhauled to stress objectives and performance and to preserve the Agency's negotiating and operating flexibility. Although Congress cannot be expected to give A.I.D. a blank check, it may be willing to accept less funding detail in the Congressional Presentation in exchange for more information on sectoral policies and expenditures, regional programs and allocations, measures of country need and performance, and project and program impacts. Within regional or sectoral categories, A.I.D. should have some discretion, subject to congressional review, to set or modify country or project funding levels. Equally important, the Congressional Presentation should systemically inform Congress on achieved and expected performance. The performance report, which should be incorporated into the Congressional Presentation, should be reasonably detailed and should provide quantitative evidence of progress made in achieving legislatively prescribed objectives. The report should also notify Congress of the performance expected from projects already underway and from those to be undertaken with funds requested in that year's budget.

Congress should curtail earmarks and give A.I.D. more flexibility in implementing approved programs. Earmarking is another old practice that takes a toll on performance. Each year, Congress prescribes in appropriations legislation, and the reports that accompany them, precisely where the bulk of foreign assistance dollars are to go. Some earmarks are by sector, others are by country or project. Their cumulative effect is to drain the Agency of much authority and initiative to allocate resources on the basis of performance.

Although it is an old practice, earmarking had modest beginnings, but both the number of earmarks and the percentage of funds locked in by congressional directive have escalated over the years. In FY 1981, for example, less than 5 percent of the Functional Development Assistance Program was earmarked. A decade later, more than 30 percent was earmarked. The Hamilton-Gilman task force found that more than 80 percent of combined Development Assistance and Economic Support Funds were earmarked in FY 1989. The percentage is probably higher today.

A.I.D.'s 1991 study of performance management concluded that it is possible to operate with earmarks while still focusing on program results. The study recommended that earmarks be first distributed to countries making the most progress, leaving any remaining funds to be parcelled out on the basis of broader performance criteria. However, while it may be technically feasible to manage with earmarks, it would be unwise to do so. If, because of earmarks, the Agency disproportionately favors

some countries, Congress is likely to respond by expanding the scope of earmarks to ensure that other countries receive their share. If this occurs, the noose will become even tighter and the Agency will have even less flexibility. Another problem is that channeling earmarks to high performers might impel the Agency to mortgage an increasing portion of new and future resources to long-term projects. Finally, and most important, earmarks send the wrong message that performance doesn't really matter because funds are allocated to those with political influence.

Administrator Roskens correctly noted in the 1991 Congressional Presentation that extensive earmarking "swamps our efforts to focus assistance on countries that have put in place policy and institutional reforms that increase their chances for broadly based growth in the short run and sustainable development over the longer run." The Africa Bureau has demonstrated that release from earmarks can lead to a coherent, focused development strategy that closely monitors the impacts on U.S. assistance and relies on performance data in allocating resources.

It is unlikely that earmarks will disappear, and even if they did, some will quickly creep back. What is important is that Congress reverse the progressive expansion of earmarking and that A.I.D. have greater flexibility in managing its resources. The rollback of earmarks should proceed in measured paces, perhaps as a pilot test in particular sectors, with the Agency required to demonstrate improved performance to win further discretion.

Transforming A.I.D. Operations

For performance management to become part of the Agency culture, it does not suffice for Congress to loosen some of the constraints on the Agency. Nothing will really change unless the Agency changes. Moreover, Congress will not act unless it is confident that A.I.D. will use its newly won flexibility prudently and that its internal controls are effective. The changes required of A.I.D. go beyond the recent reorganization and the drive for performance management. Although welcome, these initiatives take established practices and roles for granted.

A.I.D. must reexamine itself and question basic operating tenets that have been protected for decades. It must undertake truly difficult changes, beginning with relationships among central staff, bureaus, and Missions.

Senior Management must provide clear, effective policy guidance, including objectives and performance expectations, to be used by bureaus and Missions in programming actions. Decentralization is the code word for an Agency that allows bureaus and Missions an extraordinary amount of discretion in programming decisions and resource allocations. Decentralization explains why some organizational units have taken performance seriously, whereas others have not. Decentralization is a valuable asset for an Agency that operates in 90 countries under many different conditions and with a relatively small headquarters structure. But decentralization can also be an excuse for maintaining the

status quo and a barrier to change, as well as a mask that veils A.I.D.'s inability to remake itself. Decentralization makes the drive for performance dependent on the cooperation of others, with the result that headquarters' exhortations often become substitutes for more forthright action.

In government agencies, as in other large organizations, field initiative and responsiveness to local conditions are highly desirable. But A.I.D. senior management, which is responsible for policy direction and is accountable to Congress, must have the capacity to lead. It is essential that the top provide policy guidance, so that those below have a clear sense of what is required of them. Of course, A.I.D. must be decentralized, but it must also have leadership, strategic vision, and the ability to mobilize resources and change course. It must know where it is heading. As difficult as it may have been for an Agency constrained by earmarks, A.I.D. would have responded more quickly and vigorously to revolutionary developments in Eastern and Central Europe if its leaders had really called the shots.

When operations are as far flung as are A.I.D.'s, central leadership is especially needed to direct units toward common objectives and to ensure that legislative requirements and other rules are followed. It is not always an advantage that each Mission follows its own path. It is also necessary that operating units carry out the policies set by those responsible for defending the Agency's programs and budget. Left alone, many Missions will drag their feet in moving projects to completion or in evaluating completed projects. They maintain an extended pipeline that protects them against unwanted intervention by headquarters, while ensuring a continuing flow of funds and a stable portfolio of activities.

A decentralized performance measurement system can reinforce central weakness. If objectives and measures bubble up from below and are consolidated by the Policy Directorate into Agencywide measures, senior managers will be dependent on the interests and perspectives of lower levels, and the new Directorate will be consigned to performing measurement and evaluation tasks rather than carrying out the policy role intended for it by the recent reorganization.

To guard against this, the Policy Directorate must be involved in making substantive policy, even if this takes attention away from systems design and procedures. When the top leads, it defines objectives as well as the measures by which progress is assessed, and it has the capacity to redirect projects and activities, reprogram funds from old undertakings to new priorities, and take other steps that keep bureaus and Missions attuned to the Agency's agenda. This condition does not now prevail in A.I.D. and, until it does, performance will be little more than a slogan.

Projects should be consolidated into fewer activities with potential for significant, measurable impacts on development. A.I.D. maintains a portfolio of approximately 2,000 new and

ongoing projects, many of which have numerous subcomponents. The general agreement is that 2,000 projects are far too many and the Agency will be more effective if it concentrates on fewer major activities.

Having too many projects is a byproduct of legislated earmarks and excessive decentralization. The earmarks chop development aid into functional categories that are then allocated among the Missions. The Missions have an incentive to spread their resources among many projects so that they can have a backlog of approved projects along with claims on future resources and stable relationships with the host country and A.I.D./Washington. The multitude of projects also arises from a basic fact of development: poor countries have many needs.

The plethora of projects generates serious management deficiencies (including the extended pipeline discussed in the next recommendation). Projects tend to be underfunded, typically take years to complete, are difficult to evaluate, have insufficient impacts on development, suffer from staff turnover, and have inadequate internal controls. All these drawbacks erode the Agency's capacity to keep its eye on performance.

The project is A.I.D.'s means of delivering performance. Even with the current emphasis on economic and political institutions, it is a specific project that improves nutrition or health care or brings education or family planning practices to a community. If A.I.D. is to be serious about results, it must pay close attention to what the Agency is receiving from the activities funded with U.S. assistance.

To improve project definition and evaluation, A.I.D. should sharply curtail the number of ongoing and planned activities. (Others have recommended that the list of assisted countries also be pruned.) Each approved project should be of sufficient size and scope to have a critical impact on development. Every proposed project should be put to a simple test: If the project succeeds, will it make a substantial difference in the country's well-being? Projects falling below this "difference" threshold should not be funded.

A.I.D. should shorten the pipeline by deobligating slow moving projects, adopting and enforcing realistic timetables, and avoiding the premature commitment of future resources. A.I.D. starts each year with a \$9 billion pipeline of obligated but unspent funds, a total that is almost 1 1/2 times its annual appropriation. The pipeline comprises several thousand projects, many of them 5 to 10 years old. Some of the projects are so old that when the M accounts were terminated in 1991, A.I.D. had to obtain a special waiver to retain \$420 million in pre-1985 funds. The pipeline is an offshoot of the slow pace of development and peculiarities of the Federal budget process. The pipeline would be much smaller if the pace of development were quicker or if funds were appropriated on a cash (rather than obligations) basis. But the pipeline is also symptomatic of the Agency's ills, of a chronic inability to bring actions to closure, to move

forcefully when problems arise. There is nothing inherent in the Agency's charter that compels it to stand by helplessly when deadlines are missed and scarce assistance dollars unused. The delays are costly, not only because they divert funds that might be productively used elsewhere, but also because they disable A.I.D.'s capacity to manage for results. From the time a slow project is planned to its completion, the Agency will have gone through one or more planning cycles, one or more changes in its objectives and priorities, one or more four changes in the field, numerous management shifts and initiatives in Washington, and a succession of annual budget cycles. The host country will have had its own ups and downs, often including changes in political leadership and in economic conditions. As the project languishes, Mission and country representatives will go through several rounds of negotiation, in which old objectives will be redefined and new ones introduced. Long before the books are closed on the project, its initial purposes will have faded and those completing the work will have little understanding of what was originally expected. These conditions destroy the systematic consideration of performance. Only a pipeline of active projects moving on schedule from design to completion can put the focus on performance again.

Getting an "actionable" portfolio of projects is within A.I.D.'s reach. The Agency has broad authority to deobligate and reobligate funds. Since 1987, virtually all of its appropriations have been "no year if obligated," meaning that once obligated, these funds no longer have fiscal-year limitations and they do not lapse if deobligated in a future year. A.I.D. also has "section 515" authority to deobligate and reobligate certain funds after their period of availability has expired. But the Agency rarely exercises its obligation-deobligation authority. By the end of the 1991 fiscal year, A.I.D. had deobligated on 1.3 percent of its "no year if obligated" funds for the 1987-1990 fiscal years. This low percentage is due to A.I.D.'s requirement that project officers obtain the affected country's consent before deobligating funds. However, the Agency plans to insert a clause into future country agreements specifying conditions under which it can unilaterally deobligate certain funds.

Although this provision might encourage some Missions to shift funds from laggard projects to higher priority activities, the odds are that without strong central leadership not much progress can be made in redoing the backlog of unfinished projects. Field staff, for many reasons, are reluctant to end nonperforming projects, even if they have the authority to do so. Central officials may be more willing to terminate troubled projects and to reprogram any remaining funds.

5. Concluding Thoughts

The recommendations presented in this report are necessary but not sufficient for vitalizing performance management in A.I.D. One more ingredient is required: An act of will by A.I.D.'s leadership to remake the Agency. Reform is not enough, for it connotes technique and tinkering, not the painful breaking down

of old habits and practices. The analogy that comes to mind is of individuals who are "born again," not in the religious sense, but in the sense of standing apart from themselves, reviewing their behavior and character, and not liking what they see. And with true discipline and determination they transform themselves. As hard as it is for individuals to transform themselves, it is even more difficult for organizations to change, because of the many different convictions and personalities involved.

Words alone will not effect change. Exhortation must be given meaning by deeds that clearly demonstrate that performance matters, that running successful projects and meeting objectives translates into more funds and more discretion. PBB is only a tool for helping leaders to remake their organization. It will be of genuine utility only if leaders really want to use it.

FOOTNOTES:

1. Sunnyvale's version of performance budgeting is discussed in Osborne and Gaebler 1992, pp. 142 - 145.
2. Three guidebooks were published in 1950, the year after the Hoover Commission recommended performance budgeting: Techniques for Development of a Work Measurement System, 1950, Work Measurement in Performance Budgeting and Management Improvement, and A Work Measurement System.
3. The application of PPB systems in conducting foreign policy is discussed in Mosher and Harr 1970, pp. 8-13, 90-98.
4. The 1973 and 1974 objectives were discussed in two articles by Havemann 1973, pp. 783-793 and 1974, pp. 1597-1605.
5. Incremental budgeting is vigorously explained and defended in Wildavsky 1964.
6. Statement of Frank Hodsoll before the Senate Committee on Governmental Affairs, May 5, 1992.
7. More than 30 studies are summarized in the House Committee on Foreign Affairs' Foreign Assistance Policy Studies: A Review of Major Reports and Recommendations (1988). See also Development and the National Interest: U.S. Economic Assistance Into the 21st Century (1989, 155-158), and Report to the President: An Action Plan (1992, 60-61).
8. Under the revised guidance issued in 1980, Missions were instructed to address a series of questions in preparing the CDSS. These included, To what extent do macroeconomic factors inhibit or enhance the ability of the poor to raise their incomes? How responsive is the political structure to the needs of the poor? How mobile is the society? How much control is there on the free flow of information? See Agency of International Development, Airgram AID-35, October 1980.
9. These and other shortcomings are discussed in A.I.D./OMB 1992.

Bibliography

Agency for International Development. 1961. "The Reorganization of Foreign Aid of 1961." Washington, D.C.: A.I.D.
Agency for International Development. 1978. "Guidance for the

Country Development Strategy Statement (CDSS)." Airgram A-384. Washington, D.C.: A.I.D. September 16.

Agency for International Development. 1985. Blue-Print for Development. Washington, D.C.: A.I.D.

Agency for International Development. 1989. Development and the National Interest: U.S. Economic Assistance into the 21st Century. Washington, D.C.: A.I.D.

Agency for International Development. 1991. "A Report of the Subcommittee on Performance Measurement." Washington, D.C.: A.I.D.

Agency for International Development. 1991. Congressional Presentation. Washington, D.C.: A.I.D.

Agency for International Development. 1993. Congressional Presentation. Washington, D.C.: A.I.D.

Agency for International Development, Bureau for Africa. 1991. FY91 Guidance on the Assessment of Program Impact. Washington, D.C.: A.I.D. August 17.

Agency for International Development/Office of Management and Budget SWAT Team No. 5. 1992. "Project and Program Evaluation." Washington, D.C.: A.I.D.

Booz-Allen & Hamilton. 1965. The Administration of the Agency for International Development. Washington, D.C.: Booz-Allen & Hamilton. February.

The Budget of the United States Government, Fiscal Year 1975. 1974. Washington, D.C.: Government Printing Office.

Burkhead, Jesse. 1956. Government Budgeting. New York: John Wiley & Sons.

The Commission on Security and Economic Assistance (the Carlucci Commission). 1983. "A Report to the Secretary of State." Washington, D.C. Government Printing Office.

Development Alternatives. 1991. "Evaluation of Performance Measurement Systems for Use in Allocating Development Assistance." Washington, D.C.: Development Alternatives. June.

Hatry, Harry P., et al. (eds). 1990. Service Efforts and Accomplishments Reporting: Its Time Has Come: An Overview. Norwalk, CT: Government Accounting Standards Board.

Havemann, Joel. 1973. "OMB Begins Major Program to Identify and Attain Presidential Goals." National Journal. June:783-93

Havemann, Joel. 1974. "Ford Endorses 1972 Goals of Management by Objective Plan." 1974. National Journal. October: 1597-1605.

Mosher, Frederick C., and John E. Harr. 1970. Programming Systems and Foreign Affairs Leadership. New York: Oxford University Press.

Office of the City Auditor. 1992. "City of Portland Service Efforts and Accomplishments: 1990-91." Portland, Oregon. March.

Oregon Progress Board. 1991. "Oregon Benchmarks: Setting Measurable Standards of Progress." Portland, Oregon. January.

Osborne, David, and Ted Gaebler. 1992. Reinventing Government. Reading, Massachusetts: Addison-Wesley Publishing Company.

The President's Commission on the Management of A.I.D. Programs. 1992. "Report to the President: An Action Plan." Washington, D.C. April.

President's General Advisory Committee on Foreign Assistance Programs. 1968. "Development Assistance in the New Administration." Washington, D.C. October.

Program Procedures Task Force. 1977. "A.I.D. Action Memorandum

for the Administrator."

Public Management Service, Organization for Economic Cooperation and Development. 1992. "Performance Management in the Public Sector: A General Overview." Unpublished report. January.

Rose, Richard. 1976. Managing Presidential Objectives. New York: The Free Press.

Rose, Richard. 1977. "Implementation and Evaporation: The Record of MBO." Public Administration Review. 37:68.

U.S. Bureau of the Budget. 1950. Techniques for Development of a Work Measurement System. Washington, D.C.: Government Printing Office.

U.S. Bureau of the Budget. 1950. Work Measurement in Performance Budgeting and Management Improvement. Washington, D.C.: Government Printing Office.

U.S. Bureau of the Budget. 1950. A Work Measurement System. Washington, D.C.: Government Printing Office.

U.S. Commission on Organization of the Executive Branch of the Government. 1949. Budgeting and Accounting. Washington, D.C.: Government Printing Office.

U.S. Congress, House Committee on Foreign Affairs. 1988. Foreign Assistance Policy Studies: A Review of Major Reports and Recommendations. Committee Print, 10th Cong., 2d Sess. Washington, D.C.: Government Printing Office. August.

U.S. Congress, House Committee on Foreign Affairs. 1989. Report of the Task Force on Foreign Assistance to the Committee on Foreign Affairs. Washington, D.C.: Government Printing Office. February.

U.S. General Accounting Office. 1991. Foreign Assistance: Progress in Implementing Development Fund for Africa. GAO/NSIAD 91-127. Washington, D.C.: GAO. April.

U.S. General Accounting Office. 1992. Program Performance Measures: Federal Agency Collection and Use of Performance Data. Report No. GGD-92-65. Washington, D.C.: GAO. May.

U.S. Senate, Committee on Government Operations. 1977. Compendium of Materials on Zero-Base Budgeting. Washington, D.C.: Committee Print. January.

Wildavsky, Aaron. 1964. The Politics of the Budgetary Process. Boston, Massachusetts: Little, Brown & Co.

White, Louise G. 1986. An Approach to Evaluating the Impact of A.I.D. Projects. Program Design and Evaluation Methodology Report No. 5. Washington, D.C.: A.I.D. March.